



Katholieke Universiteit Leuven
Centrum voor Onderwijs sociologie

**Vlerick Leuven Gent
Management School**

Vlerick Leuven Gent
Management School
Competentiecentrum
Mens & Organisatie

Universiteit Antwerpen
Departement
Didactiek en Kritiek

Een alternatieve kijk op evaluatie

Eindverslag

april 2002

P. Van Petegem
J. Vanhoof
J. C. Verhoeven
G. Devos
V. Warmoes

Onderzoek uitgevoerd in opdracht
van het Departement Onderwijs
O.B.P.W.O.-Project 99.11

© Universiteit Antwerpen / Departement Didactiek en Kritiek
Universiteitsplein 1
2610 Wilrijk
België

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke wijze ook, zonder voorafgaande schriftelijke toestemming van de auteurs.

Inhoudsopgave

	Pagina
Ten geleide	5
Deel 1: Stilstaan bij evaluatie	17
Deel 2: Op zoek naar goede evaluatie	57
Deel 3: Aandacht voor het evaluatieproces	111
Deel 4: Alternatieve evaluatievormen geïllustreerd	173
Deel 5: Aanvangen met alternatieve evaluatievormen	269
Verklarende woordenlijst	291
Lijst met figuren en denkoefeningen	295
Literatuurlijst	300

**Ten
geleide**

INHOUD

Ten geleide

<u>1 Inleidend</u>	7
<u>2 Structuur van het handboek</u>	9
<u>3 Een leeswijzer</u>	10
<u>4 Gehanteerde termen</u>	11
<u>4.1 Evaluatie: verschillen over de taalgrenzen</u>	11
<u>4.2 Termen in het betoog over alternatieve evaluatievormen</u>	13
<u>4.3 Aanverwante evaluatietermen</u>	14
<u>5 Afsluitend</u>	16

1 Inleidend

Wat is evaluatie? Een veelheid aan invullingen van de term 'evaluatie' zou als antwoord op deze vraag naar voren kunnen worden geschoven. Het belang van evaluatie kan dan ook moeilijk overschat worden. Ze speelt een beslissende rol in de dagelijkse onderwijspraktijk, zowel ten behoeve van de leerling, de leerkracht, de school als beleidsmakers en ouders. Getuige hiervan volgende greep uit mogelijke vormen van evaluatie.

Figuur 0.1: Vormen van evaluatie

- ✓ Evaluatie kan de **leerling** duidelijkheid verschaffen over wat van hem verwacht wordt en in welke mate hij al dan niet aan de vooropgestelde criteria beantwoordt. Duidelijkheid aangaande de doelstellingen, criteria en verwachte standaarden kunnen in sterke mate de studieaanpak en motivatie bevorderen. De leerlingen worden door de evaluatie gestimuleerd om over hun eigen leerproces te reflecteren. In overleg met de leerling worden evaluatiegegevens eveneens gebruikt om beslissingen te nemen over de schoolloopbaan of de beroepskeuze van de leerling.
- ✓ Evaluatie kan de **leerkracht** informeren over de eigen instructie en de resultaten van de leerlingen. Op grond van de evaluatiegegevens kan de leerkracht beslissingen nemen omtrent de vraag of het al dan niet wenselijk is het onderwijsleerproces bij te sturen en in welke richting dit best gebeurt. Evaluatie kan de leerkracht van gegevens voorzien om leerlingen in niveaugroepen in te delen en indien nodig individueel te begeleiden. Leerkrachten gebruiken evaluatie bovendien ook om leerlingen te motiveren.
- ✓ Evaluatie is eveneens een belangrijk instrument voor de **school**. Het komt er op aan de onderwijskwaliteit in kaart te brengen, sterke en zwakke punten in het schoolbeleid te identificeren en deze te bestendigen dan wel te optimaliseren. Op basis van evaluatiegegevens kunnen prioriteiten bepaald worden en kan men op zoek gaan naar succesvolle alternatieven voor gangbare praktijken.
- ✓ **Beleidsmakers** gebruiken evaluatiegegevens om de kwaliteit van het onderwijs te bewaken, om doelstellingen te formuleren, om goed/slecht functionerende scholen te belonen/te sanctioneren, om financiële middelen vrij te maken voor scholen met specifieke leerlingengroepen, etc.
- ✓ Wanneer evaluatiegegevens van leerlingen gebruikt worden om scholen ter verantwoording te roepen en bovendien openbaar worden gemaakt, dan kunnen **ouders** die evaluatiegegevens bijvoorbeeld gebruiken om een gepaste school te kiezen voor hun kind.

Kortom: het belang van evaluatie blijkt op verschillende domeinen en niveaus (Alberts, 1990; Herman et al., 1991; Winograd & Perkins, 1996; Arrasmith et al, 1996; Standaert & Troch, 1998). Wellicht kan u zelf deze opsomming verder aanvullen en illustreren. In dit handboek richten we ons echter niet op het hele palet van onderwijsgerelateerde evaluatievormen. We focussen op de evaluatie van leerlingen en op de informatie die op basis daarvan kan verzameld worden. Daar zijn een aantal redenen voor.

Vanuit verschillende invalshoeken wordt benadrukt dat een kanteling naar een nieuwe, alternatieve benadering van onderwijskundig evalueren wenselijk is. Voor Vlaanderen heeft de intrede van de eindtermen en ontwikkelingsdoelen een belangrijke impuls gegeven om op een ruimere

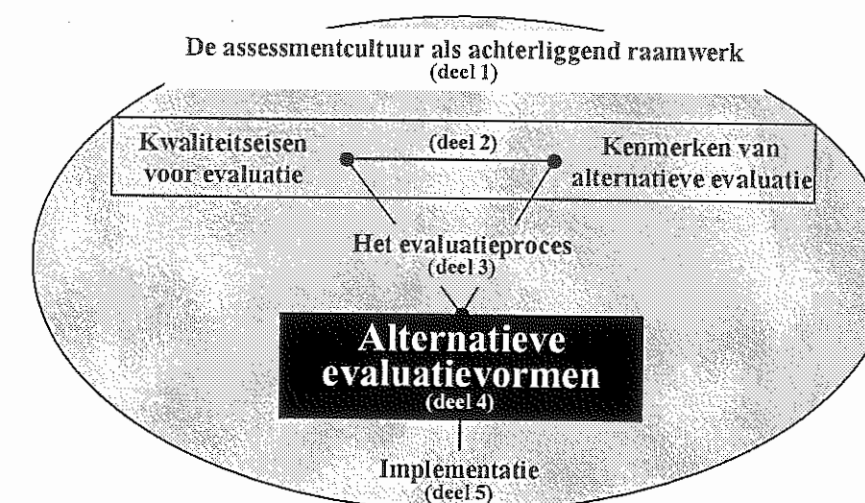
wijze naar het presteren van leerlingen te kijken. De eindtermen en ontwikkelingsdoelen omvatten immers naast doelstellingen van kennis en inzichten ook vaardigheden en attitudes. De aard van de doelstellingen is mee bepalend voor de evaluatiemethode. Nieuwe evaluatievormen worden echter niet als vanzelfsprekend geïmplementeerd door de invoering van de eindtermen (vgl. Creemers, 1991). Daarom dienen leerkrachten te worden aangemoedigd en ondersteund bij het gebruik van nieuwe beoordelingsmethoden (OECD, 1998). Dit handboek wil zich echter niet beperken tot leerkrachten. Iedereen die geïnteresseerd is in het ontwerpen van nieuwe evaluatievormen rekenen we tot onze doelgroep. Te denken valt aan praktiserende leerkrachten, leerkrachten in opleiding, begeleiders en beleids mensen op de verschillende onderwijsniveaus en onderwijsonderzoekers.

De hernieuwde aandacht die aan het evalueren van leerlingen wordt besteed, valt samen met een groeiende ontevredenheid over traditionele vormen van evaluatie. De klassieke manieren van evalueren (in hoofdzaak toetsing) doen immers onvoldoende recht aan het brede scala van onderwijsdoelen dat de overheid nastrevenswaardig acht in het Vlaamse onderwijs. Ook in de internationale literatuur stelt men vast dat het klassieke testen niet beantwoordt aan de vragen en noden van de hedendaagse samenleving. Woord en daad blijken bij evaluatie bovendien vaak uit elkaar te liggen. Er bestaat een kloof tussen onderwijskundige inzichten en de evaluatieliteratuur enerzijds en de feitelijke evaluatiepraktijken anderzijds. En ook al is de evaluatiepraktijk in Vlaanderen geëvolueerd tot op een punt waarop algemeen onderschreven wordt dat het meten van kwantificeerbare en telbare indicaties van complexe menselijke prestaties niet volstaat, toch is er nog duidelijk gebrek aan een comprehensief, systematisch en geïntegreerd kader dat leerkrachten kan ondersteunen in het ontwerpen en gebruiken van alternatieve evaluatievormen. We kunnen ons in dit handboek bijgevolg niet beperken tot het beschrijven van een greep uit bestaande alternatieve evaluatievormen zonder deze in het onderwijsgebeuren (in ruime zin) te situeren.

Alvorens deze inleiding af te ronden, moeten we één zaak duidelijk benadrukken. In een handboek zoals dit kan niet voor alle leerkrachten pasklare antwoorden gegeven worden op de vraag hoe alternatieve evaluatiebenaderingen in de concrete klaspraktijk kunnen worden gerealiseerd. -gesteld dat we dat al wenselijk zouden vinden, overigens. Eerder dan het aanreiken van 'gesneden brood' willen we noodzakelijke ingrediënten en essentiële aandachtspunten voor het ontwikkelen van goede evaluatiemethoden voorstellen. Het in verschillende verhoudingen, op verschillende manieren en tijdstippen samenbrengen van deze ingrediënten, kan leiden tot verscheidene evaluatievormen. We willen leerkrachten de nodige inzichten en strategieën aanbieden om zelf aan de slag te gaan. Dit neemt echter niet weg dat we voor een aantal leeftijdsgroepen en studierichtingen concreet voorbeeldmateriaal zullen presenteren.

2 Structuur van het handboek

Dit boek bestaat uit vijf delen. Van elk deel geven we kort de inhoud weer.



Figuur 0.2: Structuur van het handboek

- ✓ In een eerste deel presenteren we een kader waarin het ontwikkelen van alternatieve evaluatievormen vorm zal krijgen. We tonen aan dat een evaluatiemethode steeds gebaseerd moet zijn op een visie op leren. Daarbij zijn zowel de traditionele als de hedendaagse onderwijs- en evaluatiebenadering van belang. Centraal hierin staan een aantal ontwikkelingen die we als volgt zullen samenvatten: van een traditionele toetscultuur naar een assessmentcultuur. Aansluitend bestuderen we de huidige stand van zaken ten aanzien van de assessmentcultuur.
- ✓ In het tweede deel stellen we ons de vraag wat de kenmerken zijn van evaluatie in een assessmentcultuur. Het antwoord zal leiden tot een aantal kwaliteitseisen ten aanzien van goede evaluatie in het algemeen en van alternatieve evaluatievormen in het bijzonder. We ma-

ken dus expliciet het onderscheid tussen beide groepen eisen aan evaluatie; tussen goede evaluatie en alternatieve evaluatie. Daarmee wordt benadrukt dat een evaluatie niet per definitie alternatief moet zijn om als 'goed' te worden bestempeld. Alternatieve evaluatie gaat als het ware een stap verder dan goede evaluatie.

- ✓ Om te komen tot goede én alternatieve evaluatie moet zorgvuldig aandacht worden besteed aan het ontwikkelingsproces van de evaluatie. Dit proces staat centraal in het derde deel. Wanneer je als leerkracht een evaluatie op touw wil zetten, dan kunnen een aantal stappen naar voren worden geschoven die moeten doorlopen worden. Deze stappen lichten we in dit deel toe.
- ✓ Het vierde deel van dit handboek beschrijft een aantal evaluatievormen die de term alternatieve evaluatievorm verdienen. Bij elke evaluatiemethode geven we aan welke kenmerken maken dat deze benaming terecht is. Achtereenvolgens komen vaardigheidsevaluatie, authentieke evaluatie, zelf- en peerevaluatie, collaboratieve evaluatie en het gebruik van portfolio's aan bod. Het is hier dat het eigenlijke gebruik van alternatieve evaluatievormen geïllustreerd zal worden.
- ✓ In het vijfde en laatste deel stellen we het aanvangen met (of de implementatie van) alternatieve evaluatievormen centraal. We identificeren mogelijke valkuilen en presenteren een aantal kritische factoren die bijdragen tot een succesvolle implementatie.

Alvorens met de beschrijving van deze verschillende delen aan te vangen, zijn nog twee zaken van belang. In de eerste plaats presenteren we in een leeswijzer de structuurelementen die het lezen en het gebruik van dit boek beogen te organiseren. Daarna verklaren we voor alle duidelijkheid een aantal termen en definities die doorheen het handboek zullen worden gebruikt.

3 Een leeswijzer

Deze publicatie is traditioneel in die zin dat er slechts gebruik gemaakt wordt van één medium: papier. Dit neemt niet weg dat ook het gebruik van papier verschillende informatievormen en mogelijkheden biedt. In deze leeswijzer zetten we de verschillende structuurelementen voor alle duidelijkheid even op een rij.

Zoals in elk geschreven document bestaat een groot deel van dit handboek uit platte tekst. De informatie in deze tekst vormt de kern van het verhaal. Om de leesbaarheid van en de reflectie op het verhaal in de platte tekst te vergroten, wordt de tekst doorspekt met figuren, denkoefeningen en activiteiten. De figuren hebben allerlei voor zich sprekende functies. Ze verduidelijken de structuur, geven voorbeelden of reiken overzichtelijke schema's aan. Bij de denkoefeningen ligt dit anders. De functie van de denkoefeningen is de lezer uit te nodigen om te reflecteren over de gepresenteerde inhoud. Onder de titel 'Even stilstaan...' wordt de lezer een spiegel voorgehouden omtrent de eigen evaluatieopvattingen en evaluatiepraktijk. In de denkoefeningen worden vragen gesteld, maar geen antwoorden gegeven. Het is net aan de lezer om op zoek te gaan naar antwoorden. Dit mag op het eerste zicht belastend lijken, toch is voldoende

aandacht voor de denkoefeningen een noodzaak om de mogelijkheden van het handboek maximaal te benutten. Een enkele keer vindt u in het boek ook activiteiten. Deze bieden gebruiksklare werkwijzen om bepaalde evaluatieaspecten en -problemen in de eigen context aan te kaarten. De figuren, denkoefeningen en activiteiten worden qua vormgeving duidelijk onderscheiden van de platte tekst.

Zoals u reeds kon merken wordt in het boek gebruik gemaakt van *vet* tekst en *schuin gedrukte tekst*. Deze ondersteunen de structuur en de leesbaarheid. Onderlijnde tekst daarentegen betekent dat bepaalde termen of ideeën van cruciaal belang zijn in het verhaal van alternatieve evaluatievormen.

4 Gehanteerde termen

Om dezelfde taal te kunnen spreken, moeten we duidelijkheid verschaffen over de gehanteerde termen. Zoals steeds zijn definities eerder saai, maar noodzakelijk. In dit inleidende deel van het handboek staan we noodgedwongen stil bij de betekenis van drie groepen van termen. Deze zijn de term evaluatie op zich, termen in verband met alternatieve evaluatievormen en een aantal andere belangrijke evaluatietermen.

4.1 Evaluatie: verschillen over de taalgrenzen

De term evaluatie kent een enigszins verschillende inhoud naargelang hij gehanteerd wordt in het Engels-Amerikaanse dan wel in het Nederlandse taalgebied.

1. Het Engels-Amerikaanse woordgebruik

In het Engelse taalgebied maakt men een begripsmatig onderscheid tussen 'evaluation' en 'assessment'. De laatste term doelt op evaluatie van personen (bijvoorbeeld leerlingen) en de eerste term op de evaluatie van objecten (curricula, projecten, instructiemateriaal, handboeken en dergelijke). Evaluatie wordt omschreven als het beslissingsproces omtrent de mate waarin opvoedkundige doelstellingen tot op heden zijn gerealiseerd door het curriculum of de instructie. De klemtoon ligt hierbij voornamelijk op de resultaten van programma's en projecten. Assessment daarentegen is het proces dat erop gericht is te ontdekken wat een persoon kent en kan. Bij assessment is men voornamelijk bezig met de individuele persoon, de student of de leerling (Harris & Bell, 1990; Madaus & Kellaghan, 1992).

2. Het Nederlandstalige woordgebruik

In het Nederlands zijn geen begrippen beschikbaar om het bovenstaande onderscheid taalkundig te maken. Vandaar dat wij ervoor opteren -waarnodig- de Engelse term 'assessment' over te

nemen en dit in combinatie met de term 'evaluatie'. Het onderscheid tussen assessment en evaluatie houden wij echter niet voortdurend aan. Daarvoor is de term evaluatie in Vlaanderen te zeer ingeburgerd als het evalueren van leerlingen. Dit betekent dat in het Nederlandstalige woordgebruik evaluatie dus wel degelijk betrekking kan hebben op het kennen en kunnen van leerlingen. Het onderwerp van een assessment kan echter niet bestaan uit een schoolplan. Evaluatie is dus ruimer dan assessment. Of omgekeerd: assessment is een deel van evaluatie.

Definities van 'evaluatie' worden in de Nederlandstalige literatuur wel gevonden. Hieronder geven we illustratief een aantal relevante definities weer.

Evaluatie is een systematische activiteit voor het verkrijgen van zo objectief mogelijke informatie, op grond waarvan uitspraken over de waarde van het evaluatieobject (een toets, een leerling, een leerplan of gedeelte daarvan, een onderwijsmethode, enz.) worden gedaan en waarover eventueel beslissingen kunnen worden genomen (Thio, 1988).

Evaluatie is het systematisch informatie verzamelen over de aard en de kwaliteit van onderwijsobjecten. Het gaat hierbij om het beschrijven en het beoordelen van deze objecten. In principe kan alles het object zijn van evaluatie, zowel de leerlingen als programma's, projecten en instructiemateriaal. Het beschrijven kan zowel kwantitatief als kwalitatief zijn. Kwantitatieve beschrijving noemen we meting. Men gaat tijdens een evaluatie aan gedragingen en onderwijsleerresultaten een waarde toekennen en op basis daarvan beslissingen nemen (Nevo, 1995).

Evalueren in de dagelijkse onderwijspraktijk is een proces waarbij de leerkracht informatie verzamelt betreffende de leerprestaties van de leerlingen, zodat hij op grond hiervan een waardeoordeel kan toekennen. In de meeste gevallen wordt daaraan een beslissing gekoppeld die uitmondt in een rapportering aan de betrokkene(n) (Standaert & Troch, 1998).

De verscheidenheid aan definities in het Vlaamse en Angelsaksisch taalgebied is opvallend. Vaak is de invalshoek verschillend en worden andere elementen benadrukt. Op basis van de bestaande definities zullen wij in dit boek dan ook een specifieke definitie hanteren. De evaluatie van leerlingen zal doorheen het handboek gedefinieerd worden als:

"Evaluatie is het verzamelen en beoordelen van gegevens over de (aan de doelstellingen van het onderwijs gerelateerde) prestaties van leerlingen."

In deze definitie komen de sleutelementen van evaluatie aan bod. Onder evaluatie verstaan we dus een tweeledig proces. Het verzamelen van informatie over de leerprestaties van leerlingen en het beoordelen van deze informatie. Daarnaast wordt de gerichtheid op de doelstellingen van het onderwijs benadrukt. Verder is de definitie doelbewust beperkt opgevat. Er wordt bijvoorbeeld niet aangegeven wat de inhoud is van de evaluatie en welke de mogelijke functies ervan zijn. Bovendien wordt in het midden gelaten door wie de gegevens verzameld dienen te worden. Dergelijke elementen kunnen niet allemaal vastgelegd worden in een werkbare definitie. De

manier waarop bovenstaande definitie verder ingevuld dient te worden, verschilt immers naar gelang van de specifieke evaluatiecontext. Tenzij ze beperkt wordt opgevat, zou de definitie van evaluatie dus telkens moeten herschreven worden.

In het verdere verloop van dit handboek zullen we uitvoerig stilstaan bij verschillende relevante aspecten van evalueren. Dit kunnen we echter niet zonder een ruimer gemeenschappelijk begrippenkader op te stellen. Volgende twee punten richten zich daarom respectievelijk op termen die gebruikt worden in het verhaal van de alternatieve evaluatievormen en op een aantal specifieke evaluatietermen.

4.2 Termen in het betoog over alternatieve evaluatievormen

Bij het voorstellen van een aantal aan alternatieve evaluatievormen gerelateerde termen is het zinvol een onderscheid te maken tussen het gebruik van deze termen in bestaande evaluatieliteratuur enerzijds en in dit handboek anderzijds.

1. Termen in bestaande evaluatieliteratuur

Om de alternatieven voor traditionele evaluatiemethoden te benoemen, worden in de onderwijskundige literatuur verschillende benamingen gehanteerd. Te denken valt aan 'alternatieve evaluatie', 'authentieke evaluatie' en 'performance-based assessment' (of kortweg performance assessment). Nuanceverschillen, verschillende invalshoeken en accentverschillen maken het echter moeilijk om een duidelijke lijn te onderkennen in het gebruik van deze verschillende termen. Dit blijkt uit de volgende figuur.

Figuur 0.3: Onduidelijkheid over begrippen

- ✓ Fisher & King (1995) spreken van **alternatieve evaluatie** wanneer de evaluatie bestaat uit een andere vorm dan de traditionele meerkeuze- en gestandaardiseerde toetsen. Anderen stellen dat alternatieve evaluatie alle types evaluatie omvat waarin leerlingen een antwoord moeten creëren in plaats van een antwoord te kiezen uit een gegeven lijst (Arter & Bond, 1996). *Bedoelen beiden hetzelfde of zijn er nuanceverschillen?*
- ✓ Men spreekt van **performance assessment** indien van leerlingen verwacht wordt dat ze een antwoord of een product creëren dat hun kennis en vaardigheden kan aantonen. Dit in tegenstelling tot traditionele evaluatievormen waar leerlingen vaak een antwoord moeten kiezen uit verschillende antwoordmogelijkheden. *Zijn alternatieve evaluatie en performance assessment synoniemen? Waarom worden er dan verschillende termen gebruikt?*
- ✓ **Authentieke evaluatie** kenmerkt zich dan weer door het gebruik van realistische contexten om kennis en vaardigheden te meten en kan dus eveneens worden afgezet tegen traditionele toetsvormen (Winograd & Perkins, 1996). Herman et al. (1992) schrijven dat zowel alternatieve als authentieke evaluatie vormen zijn van 'performance-based assessment'. Elders lezen we dat authentieke evaluatie een uitgewerkte vorm is van performance assessment. *Ik ben niet meer mee. En U?*

Nog meer dan de termen 'performance assessment' en 'authentieke evaluatie' heeft de term 'alternatieve evaluatie' doorheen de evaluatieliteratuur niet altijd dezelfde betekenis. Om die reden is het belangrijk hier reeds te omschrijven wat er in dit handboek onder zal verstaan worden. De betekenis van de andere termen in het verhaal van de alternatieve evaluatievormen brengen we later aan bod.

2. Alternatieve evaluatie in dit handboek

Op de keper beschouwd is 'alternatieve evaluatie' wat er in de term vervat zit, namelijk een alternatief. In concreto een alternatief voor de traditionele evaluatiemethoden waarvan we zullen aantonen dat ze veelal onvoldoende geschikt zijn om hedendaagse onderwijsdoelen te evalueren enerzijds en om deze hedendaagse evaluatiedoelen te realiseren anderzijds. De term 'alternatieve evaluatievormen' bevat verscheidene evaluatiemethodes en is dan ook een verzamelnaam.

In het tweede deel van dit boek beschrijven we dat een evaluatiemethode als alternatief kan beschouwd worden indien ze in voldoende mate aan een aantal eisen voldoet. Dit impliceert dat de term 'alternatief' geen kwestie is van alles of niets. Ook indien een evaluatiemethode niet aan alle vooropgestelde eisen voldoet, zullen we spreken van alternatieve evaluatie. Dit maakt 'alternatief' tot een omvattend en gradueel begrip. **Omvattend** omdat ook evaluatievormen die niet aan al de vooropgestelde eisen voldoen maar daarentegen expliciet één kenmerk benadrukken als 'alternatief' zullen bestempeld worden (te denken valt aan evaluatievormen zoals vaardigheidsevaluatie en zelfevaluatie). **Gradueel** omdat het toekennen van het begrip 'alternatief' afhankelijk is van de vraag in welke mate een bepaalde evaluatievorm tegemoet komt aan een aantal kenmerken die zelf als een continuüm kunnen worden voorgesteld.

De betekenis van 'alternatieve evaluatie' zal uiteraard verderop uitvoerig worden toegelicht. Eigenlijk volstaat om bij het afsluiten van deze inleiding het volgende te onthouden: 'alternatieve evaluatie' vormt een alternatief. Waarvoor en waaruit dit alternatief bestaat, staat centraal in de verschillende delen van dit handboek. Gaandeweg zal duidelijk worden wat we verstaan onder alternatieve evaluatie.

4.3 Aanverwante evaluatietermen

Om de aard, de functie, het tijdstip, de doelen en de kenmerken van evaluatie te omschrijven worden in de onderwijsliteratuur en -praktijk verscheidene termen gebruikt. Volgende opsomming geeft een beschrijving van de meest gehanteerde termen (Standaert & Troch, 1998; Verloop & van de Schoot, 1995; Janssens et al., 2000). Een aantal van de relevante termen bestaat uit samenhangende begrippenparen.

1. Formele en informele evaluatie

Dit eerste onderscheid betreft de te volgen evaluatieprocedure. Bij **formele** evaluatieprocedures zijn precieze voorschriften voor uitvoering van de evaluatie beschikbaar. Er is met andere woorden duidelijk omschreven waarop de evaluatie zich toelegt, welke de scoringsrichtlijnen zijn, wat de gevolgen zullen zijn van een bepaalde beoordeling en dergelijke. Bij **informele** evaluatie ontbreken dergelijke voorschriften en afspraken. Vaak doen leerkrachten een beroep op hun eigen aanvoelen zonder procedures te omschrijven. Dit is bijvoorbeeld het geval wanneer de leerkracht in de les een korte vraag stelt om na te gaan of een voorafgaande opmerking begrepen is of wanneer men bij de aanvang van de les naar de voorkennis van leerlingen peilt.

2. Formatieve en summatieve evaluatie

Het onderscheid tussen beide begrippen heeft vooral betrekking op het gebruik dat men wil maken van de evaluatiegegevens. Bij **summatieve evaluatie** wil men een eindoordeel uitspreken over de prestaties van een bepaalde leerling. Deze vorm van evaluatie wil dus als het ware een samenvatting geven van wat de leerling reeds bereikt heeft. Summatieve evaluatie is gericht op resultaatbepaling. Meestal gebeurt dit met het al dan niet toelaten van leerlingen als doel voor ogen. Bij **formatieve evaluatie** is het belang van de informatie vooral gelegen in de mogelijkheden tot feedback. Met dit evaluatietype wordt een tussentijdse vorm van evaluatie bedoeld, die de basis vormt voor de optimalisering en bijsturing van het onderwijsleerproces: enerzijds het leerproces van de leerlingen en anderzijds het onderwijzen door de leerkracht. Formatieve evaluatie wordt dus meer beschouwd als een middel tot leren dan als een maatregel tot toelating of selectie. De term 'formatief' wijst erop dat de kennis en vaardigheden waarover de leerlingen aan het einde van het leerproces moeten beschikken nog moeten 'gevormd' worden. Het onderscheid tussen formatieve en summatieve evaluatie reikt duidelijk verder dan het tijdstip van de evaluatie. Het verschil ligt vooral in de bedoeling van de evaluatie en het effect ervan op de leerlingen.

3. Proces- en productevaluatie

Het onderscheid tussen proces- en productevaluatie berust voornamelijk op de vraag wat geëvalueerd wordt. Bij **productevaluatie** wordt nagegaan in welke mate leerlingen de onderwijsdoelen bereikt hebben. Er is sprake van een rendementsevaluatie met als object van evaluatie de doelstellingen. Op deze manier evalueert men onrechtstreeks de kwaliteit van het onderwijs(gedrag) van de leraar, omdat men het via leerlingprestaties doet. Als men de verkregen resultaten gebruikt om de prestaties van leerlingen of het onderwijsgedrag van de leerkracht te verbeteren, dan krijgt deze productevaluatie een formatief karakter. Gebruikt men deze resultaten om andere beslissingen te nemen (zoals selectie, overgang, oriëntatie), dan wordt de productevaluatie summatief van aard. Men kan de kwaliteit van het onderwijsgedrag evenwel ook rechtstreeks onderzoeken (en dus niet alleen via de leerlingresultaten). Dan gaat het over **procesevaluatie**. Procesevaluatie richt zich op de kwaliteit van het leerproces zelf, dat wil zeggen dat geprobeerd wordt om het leerproces van de leerling en het 'aanleerproces' (het didactisch handelen) in kaart te brengen.

4. Permanente evaluatie

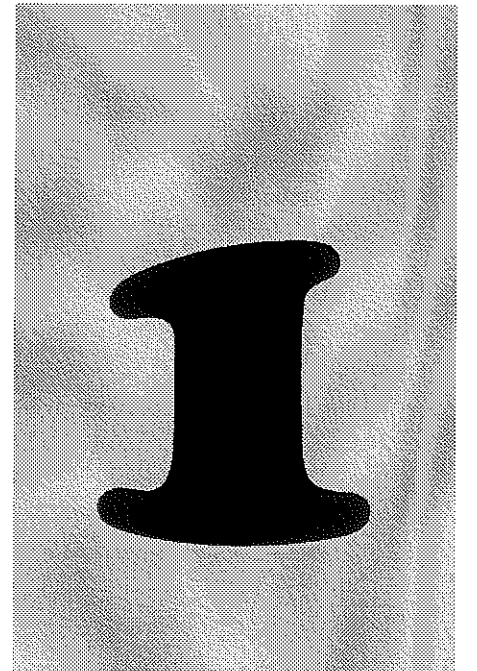
Permanente evaluatie bestaat uit een voortdurende opvolging van de vorderingen van leerlingen aan de hand van een opvolgsysteem (zoals een logboek, een leerlingvolgsysteem en dergelijke). Onder permanente evaluatie wordt dus een regelmatige evaluatie van de studieprestaties van de leerling verstaan. Tijdens het onderwijsleerproces zal de leerkracht voortdurend, door middel van doelgerichte vragen, peilen naar het rendement van zijn didactisch handelen. Tegelijkertijd zal hij het leergedrag van zijn leerlingen observeren. Zo kan hij ogenblikkelijk zijn didactisch handelen tijdens een onderwijsleeractiviteit aanpassen aan de noden van de leerlingen en leerlingen gerichte begeleiding geven.

De belangrijkste termen die in dit handboek aan bod komen, zijn achteraan samengebracht in een verklarende woordenlijst (zie pagina 291).

5 Afsluitend

Dit handboek richt zich op één specifiek aspect van evaluatie, namelijk de prestaties van leerlingen. We omschreven evaluatie dan ook als het verzamelen en beoordelen van gegevens over de (aan de doelstellingen van het onderwijs gerelateerde) leerprestaties van leerlingen. In deel 2 stellen we ons de vraag hoe dit op een kwaliteitsvolle én alternatieve manier kan gebeuren. Het is echter noodzakelijk eerst uitvoeriger stil te staan bij een aantal ruimere aspecten van evaluatie. Dit doen we in deel 1.

Deel



Stilstaan bij evaluatie

INHOUD - DEEL 1

Stilstaan bij evaluatie

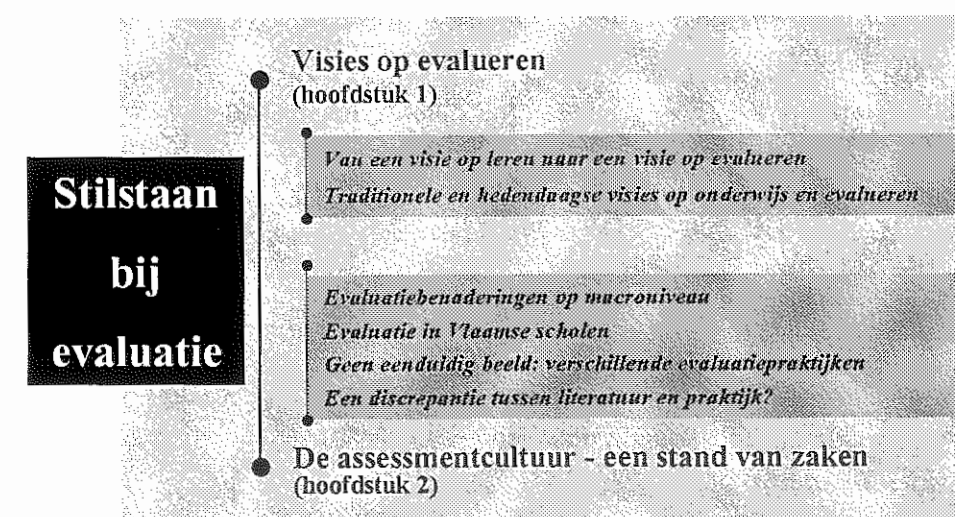
Inleidend.....	19..
Hoofdstuk 1: Visies op evalueren.....	20
1 <i>Waarom moet een evaluatiemethode gebaseerd zijn op een visie op leren?</i>	20
1.1 Van een visie op leren naar een visie op evalueren.....	20
2 <i>Traditionele en hedendaagse visies op onderwijs en evalueren</i>	21
2.1 Evoluerende onderwijsbenaderingen.....	22
2.2 Evoluerende evaluatiebenaderingen: van toetspraktijk naar een veelzijdige, geïntegreerde assessmentcultuur.....	30
2.3 Traditioneel en alternatief: tegengesteld of complementair?.....	40
Hoofdstuk 2: Assessmentcultuur - een stand van zaken.....	43
1 <i>Evaluatiebenaderingen op macroniveau</i>	43
1.1 Eindtermen en ontwikkelingsdoelen.....	43
1.2 Evalueren: regelgeving en kwaliteitsbewaking.....	45
1.3 Basiscompetenties van de leerkracht.....	46
1.4 Afrondend.....	48
2 <i>Evaluatie in Vlaamse scholen</i>	48
2.1 Kenmerken van de Vlaamse evaluatiepraktijk.....	48
2.2 Concluderend: bestaat de assessmentcultuur vooralsnog enkel op papier?.....	51
3 <i>Geen eenduidig beeld: verschillende evaluatiepraktijken</i>	52
3.1 Beschrijving van te onderscheiden evaluatiepraktijken.....	52
3.2 Concluderend: bestaat de assessmentcultuur enkel op papier?.....	53
4 <i>Een discrepantie tussen literatuur en praktijk?</i>	54
4.1 Indicaties voor de discrepantie.....	54
4.2 De discrepantie verklaard.....	55

Inleidend

Evaluatie verandert omdat onderwijs verandert. Wat de maatschappij van leerlingen verwacht, is echter niet steeds gelijk aan wat op school wordt geleerd. Om de snel veranderende maatschappij bij te benen, moet de school mee evolueren wil ze (in alle opzichten) succesvolle leerlingen opleiden en opvoeden. Veranderingen in evaluatie kunnen deze noodzakelijke veranderingen in curriculum en in instructie stimuleren en ondersteunen. Meer nog, we zullen stellen dat zonder aangepaste evaluatiemethoden goed bedoelde veranderingen in curriculum en instructie gedoemd zijn om aan hun doel voorbij te gaan. Het loont dus zeker de moeite om uitvoerig aandacht te besteden aan evaluatie in het algemeen en aan alternatieve evaluatievormen in het bijzonder.

Voorafgaand aan het beschrijven van alternatieve evaluatievormen en het ontwikkelen van concreet voorbeeldmateriaal dienaangaande, is het aangewezen aandacht te besteden aan een aantal belangrijke elementen die een algemeen kader vormen voor de verdere zoektocht naar goede evaluatievormen. Stilstaan bij evaluatie dus.

Figuur 1.1: Structuur van deel 1



In een eerste hoofdstuk verdedigen we de stelling dat een evaluatiemethode steeds moet gebaseerd zijn op een visie op leren. Vervolgens beschrijven we de traditionele en de hedendaagse onderwijs- en evaluatiebenadering. Een kijk op evalueren die we als volgt kunnen samenvatten: van een traditionele toetscultuur naar een assessmentcultuur (Struyf, 1999; Birenbaum, 1995). In het tweede hoofdstuk bestuderen we de huidige stand van zaken in Vlaanderen ten aanzien van de vooropgestelde assessmentcultuur.

Hoofdstuk 1: Visies op evalueren

Bij het kiezen van een geschikte evaluatiemethode is de uitgangsvraag die we ons steeds moeten stellen: wat willen we evalueren? Om deze vraag te beantwoorden, moeten we de doelstellingen en de context van het beoogde leerproces in overweging nemen (Birenbaum, 1995). Dit betekent dat het ontwikkelen van een visie op evalueren tevens een duidelijke visie op verscheidene aspecten van leren en onderwijzen vereist. In de volgende paragrafen wordt verduidelijkt dat deze gedachtegang ook aan de basis ligt van de voorliggende beschrijving en ontwikkeling van alternatieve evaluatievormen. Eerst belichten we kort de samenhang tussen visies op leren en visies op evalueren. Daarna staan we stil bij de manier waarop de onderwijs- en de evaluatievisies traditioneel en in de hedendaagse onderwijsliteratuur worden ingevuld.

1 Waarom moet een evaluatiemethode gebaseerd zijn op een visie op leren?

Om de samenhang tussen een visie op onderwijs, het curriculum, het onderwijsleerproces en de evaluatie kort te illustreren, doen we een beroep op een mogelijke onderwijsopvatting zoals deze door Kulieke et al. (1990) beschreven wordt (andere onderwijsopvattingen zouden een ander maar gelijkaardig verhaal opleveren).

1.1 Van een visie op leren naar een visie op evalueren



Kulieke et al. (1990) beschrijven een visie op leren die is gebaseerd op cognitieve, filosofische en multiculturele perspectieven. Dergelijke perspectieven suggereren dat men pas van zinvol leren kan spreken als de lerende beschikt over een kennisbasis die vlot kan worden aangewend om betekenis te geven aan de wereld, om problemen op te lossen en om beslissingen te nemen. De lerende moet zijn leerproces onder controle hebben en continu streven naar het verwerven en gebruiken van nieuwe technieken. Er moet tevens sprake zijn van strategisch leren met behulp van een repertoire van effectieve leerstrategieën. De lerende moet tenslotte empathisch zijn en zichzelf en de wereld vanuit verschillende (i.e. andere dan de eigen) perspectieven benaderen.

Als het doel van leren in scholen op die manier wordt gedefinieerd, moeten zowel het curriculum als de klasomgeving herdacht worden. Het curriculum (c.q. de leerinhoud) moet diepteverwerking stimuleren en leerlingen aan de hand van relevante taken in een betekenisvolle, realistische leeromgeving onderbrengen. Lerenden moeten zich inlaten met holistische leertaken en steeds gestimuleerd worden hun voorkennis te gebruiken. Een dergelijk curriculum uit zich in wat Kulieke et al. (1990) de 'collaboratieve klas' noemen. Deze wordt gekenmerkt door het delen van kennis en autoriteit door leerlingen en

leerkrachten. De leerkrachten hebben een mediërende rol en er worden heterogene klasgroepen samengesteld.

Deze veranderingen in de visie op leren, het curriculum en de onderwijspraktijk leiden eveneens tot veranderingen in het evalueren van leerlingen. Wanneer de onderwijservaringen van leerlingen veranderen, moeten de manieren die voorheen gehanteerd werden om het leren van leerlingen te evalueren immers mee evolueren om hiaraan tegemoet te komen. Stelt u zich eens even voor hoe twijfelachtig het succes zou zijn van de voorgestelde visies op leren en instructie indien leerlingen op het einde van de rit beoordeeld zouden worden aan de hand van kennisvragen die de reproductie van de geleerde stof centraal stellen. Ook de visie op evalueren als dusdanig heeft dan ook een lange ontwikkelingsweg afgelegd.

In welke richting evaluatie evolueert, vormt de kern van dit boek. We zullen op zoek gaan naar alternatieven, maar niet zonder ons bewust te zijn van de samenhang tussen onderwijsdoelen, instructie en evaluatie. Het afzetten van de hedendaagse visies tegen de eerder traditionele visies is in dit opzicht verhelderend. Eerst even een korte denkoefening.

Even stilstaan...1.1: Samenhang tussen doelen, curriculum en evaluatie

Haal een vertrouwde en concrete onderwijscontext voor de geest (die je zelf als leerkracht verzorgt of die je als student ondergaat of onderging) en achterhaal wat volgens u de belangrijkste doelstelling is of was van die onderwijssituatie. Schrijf dit zonodig even op. Met dit doel voor ogen richten we ons op de instructie en de evaluatie. De twee vragen hierbij luiden als volgt:

- ✓ Is/was de gebruikte evaluatiemethode geschikt om dat specifieke doel te evalueren? Waarom wel/niet?
- ✓ Draagt/droeg de evaluatiemethode bij tot het bereiken van de doelstelling. Op welke manier (net niet)?

2 Traditionele en hedendaagse visies op onderwijs en evalueren

De uitgangsvraag om evaluatie vorm te geven, is fundamenteel pedagogisch van aard, namelijk: welke onderwijsdoelstellingen moet men vooropstellen, wil men leerlingen voorbereiden om binnen en buiten de school succesvol te functioneren in het heden en in de toekomst. Het antwoord op deze vraag varieert naargelang de specifieke tijdsperiode waarin het onderwijs en leren plaatsvindt. Hieronder worden illustratief de traditionele en hedendaagse onderwijs- en evaluatieopvattingen naast elkaar geplaatst. Inzicht in de onderwijsopvattingen kan een aanzet geven tot verklaring van de veranderende visies op evalueren doorheen de tijd. Uit de evoluerende onderwijsbenaderingen zal immers blijken dat de evolutie in evaluatie (van eertoets- naar een assessmentcultuur) mede het gevolg is van de veranderde doelen en verwachtingen die men ten aanzien van het onderwijs vooropstelt.

2.1 Evoluerende onderwijsbenaderingen

Een aantal theoretische stromingen in de psychologie hebben geleid tot vernieuwde opvattingen over leren in onderwijssituaties. Ook de evolutie van de traditionele naar de hedendaagse onderwijsbenaderingen is mede terug te brengen tot evoluties in leertheorieën. In het onderstaande komen de traditionele en hedendaagse onderwijsbenaderingen met de bijhorende leertheorieën kort aan bod (Verschaffel & Vermunt, 1998; Herman et al., 1992; Fisher & King, 1995; Struyf et al., 1999; Arter & Bond, 1996; Imbrecht & Van Petegem, 2001; Janssens et al., 2000).

2.1.1 Traditionele onderwijsbenaderingen

Birenbaum (1995) beschrijft scholen in de tweede helft van de 19^{de} eeuw als socialisatie-instellingen die de leden van de samenleving wilden voorzien van de noodzakelijke vaardigheden om in de industriële maatschappij te kunnen functioneren. Een geschoold iemand moest de taken waarvoor hij werd opgeleid naar behoren kunnen uitvoeren. Deze vaardigheden en de informatie die ze vereisten, werden als vastliggend en eindig beschouwd. Het onderwijs voor de massa benadrukte het verwerven van basale vaardigheden terwijl hogere denkvaardigheden en intellectuele bezigheden gereserveerd waren voor een kleine elite.

De traditionele onderwijsbenadering is –voornamelijk in de Verenigde Staten– gebaseerd op een behavioristisch leermodel. Dat model beschouwt de leerling als iemand die passief de gegevens en informatie opneemt die door de leerkracht worden aangeboden. Leren wordt beschouwd als het vullen van een leeg vat. De didactische kenmerken van de behavioristisch georiënteerde opleiding zijn onder andere dat de instructie gefragmenteerd is in kleine stappen, dat ze gestuurd wordt door de leerkracht en dat die voortdurend controleert of een beoogd resultaat wordt bereikt. Vertalen we dit naar de instructie, dan wil dat zeggen dat men probeert om de complexe problemen die men wil aanleren, opbreekt in kleinere stukjes. Elke stap steunt daarbij op de kennis en kunde die in vorige stap(pen) werden geleerd. Dit leermodel heeft ertoe geleid dat leerkrachten hun toevlucht hebben genomen tot 'transmission teaching': het louter overdragen van kennis (Struyf, 1999). In de traditionele onderwijsopvatting staat de overdracht van kennis door de leerkracht centraal. Die doceert als expert in de leerstof op een overzichtelijke, op voorhand vastgelegde manier. De leerling is receptief, hij luistert en noteert wat de docent vertelt. Volgens de traditionele onderwijsbenadering bestaat leren uit het opnemen vankant-en-klare kennis en procedures die door de vorige generaties ontdekt werden. De nadruk ligt op cognitieve aspecten en de evaluatie richt zich op het leerproduct, namelijk de verworven kennis.

2.1.2 Hedendaagse onderwijsbenaderingen

Als gevolg van een technologische revolutie bevinden we ons momenteel in het zogenaamde informatietijdperk. Nieuw is het feit dat informatie niet meer aangezien wordt als statisch en eindig, maar als dynamisch en infinitief. Er is teveel kennis om ze allemaal over te dragen en

paraat te houden bij de leerlingen. Bovendien is de school niet meer de enige bron van kennis voor de leerling. Het traditionele onderwijs wordt door deze ontwikkelingen steeds meer als ontoereikend ervaren. Er worden ten aanzien van het onderwijs andere eisen gesteld (Gipps, 1994). Een heroriëntatie was dan ook noodzakelijk. In de laatste helft van de 20ste eeuw brak geleidelijk aan een nieuw psychologisch leermodel door: leren is een activiteit van het kind zelf dat handelend en in gesprek met anderen zijn/haar cognitieve (kennis)structuren bouwt. Wat kenmerkt, afgezet tegen bovenstaande traditionele onderwijsopvatting, dan de bijbehorende onderwijsdoelen en het bijbehorende onderwijsdenken?

Indicatief voor de hedendaagse onderwijsdoelen is volgende opsomming van een aantal vaardigheden die werkgevers volgens Carnevale et al. (1996) verwachten van toekomstige werknemers: leervaardigheden (weten hoe men best kan leren), lees-, schrijf- en rekenvaardigheden, communicatieve vaardigheden (zoals luisteren en mondeling communiceren), creatief denken en probleemoplossende vaardigheden, persoonlijk management (zoals zelfwaardegevoel, doelmatigheidsbeleving, motivatie), effectief in groep kunnen werken (interpersoonlijke vaardigheden, onderhandelen, teamwork) en kunnen omgaan met invloed (leiderschap en het begrijpen van functionele structuren).

Niet alleen de onderwijsdoelen maar ook het leren zelf evolueert. Op basis van de ontwikkeling van nieuwe leertheorieën is er veel meer inzicht ontstaan in de complexiteit van het leerproces. Leerprocessen kennen geen standaard; iedereen leert op zijn of haar eigen wijze. Dat maakt het aanbieden van onderwijs er niet eenvoudiger op. Onderwijs van nu moet daarom gericht zijn op voortdurende verandering, op zelfontplooiing en actief bezig zijn. Kennisoverdracht vervaagt en leren wordt steeds meer het aanreiken en ontdekken van vaardigheden om zelf kennis te construeren. Leren wordt niet meer aangezien als een éénrichtingsverkeer van de leerkracht naar de passief ontvangende leerlingen. Betekenisvolle instructie vereist een actieve bijdrage van de leerlingen aan het leerproces (Herman et al, 1992). Er ligt minder nadruk op het overbrengen van vaststaande feiten en overgeërfde kennis, maar meer op interpretatieschema's voor de complexiteit die ons omringt.

Een aantal aspecten van de hierbij aansluitende, hedendaagse onderwijsbenadering worden in de onderstaande paragrafen toegelicht. Achtereenvolgens komen het gebruik van het sociaalconstructivisme als cognitieve leertheorie, het curriculum, het onderwijsleerproces en de evaluatie aan bod.

1. Sociaalconstructivisme als cognitieve leertheorie

In tegenstelling tot de traditionele onderwijsbenadering, die gebaseerd is op behavioristische leertheorieën, zijn de hedendaagse onderwijsbenaderingen gebaseerd op de cognitieve leertheorie en het sociaalconstructivistische gedachtegoed (Birenbaum, 1995; Herman, 1992). De cognitieve psychologie ziet leren als een informatieverwerkingsproces; ze gaat na hoe leerlingen informatie verwerken en hoe die informatie in het geheugen wordt opgeslagen. Een zeer recente stroming binnen deze cognitieve psychologie is het sociaalconstructivisme. De kernidee achter

deze theorie is dat leren een actief en constructief proces is, zelfs in die mate dat leren alleen maar goed mogelijk is als de leerlingen zelf actief construerend met de leerstof bezig zijn. De docent kan slechts de omstandigheden scheppen waarin de leerling leert. Belangrijke lessen uit deze theorieën zijn dat men leren kan leren en dat er geen 'beste' manier bestaat om iedereen te onderwijzen. Het leren vindt plaats in de lerende die nieuwe kennis en vaardigheden opbouwt, gebruik makend van reeds aanwezige kennischema's. Leren is dus een voortdurend proces waarin leerlingen continu informatie ontvangen, interpreteren en in verband brengen met wat ze reeds weten en ervaren hebben (voorkennis). Op basis van nieuwe informatie worden de eigen concepten van de wereld (mentale schema's) aangepast. De kunst van het leren is dat nieuwe informatie gekoppeld wordt aan bestaande voorkennis. Pas als de eigen concepten zijn veranderd, is er wat bijgeleerd. De leerlingen construeren dus kennis en die constructie verschilt van lerende tot lerende (Clemens et al., 1999)

Inzicht in de belangrijkste elementen van de sociaal-constructivistische leertheorieën is noodzakelijk om te kunnen meedenken over het waarom van een zoektocht naar alternatieve evaluatievormen. Vandaar dat we kort stilstaan bij een aantal sleutelkenmerken van deze leertheorieën, namelijk: leren is een actief en constructief, cumulatief, zelfregulatief, doelgericht, contextgebonden en sociaal proces.

a. Leren is actief en constructief

Betekenisvol leren is actief en constructief. Leerlingen nemen niet zozeer feitelijke informatie op, ze creëren hun eigen begrippen en structuren van de werkelijkheid. Iets weten is niet zomaar passief informatie ontvangen, maar het interpreteren en incorporeren van de informatie in de beschikbare voorkennis. In het curriculum moet men dan ook aandacht hebben voor het organiseren, structureren en gebruiken van informatie met het oog op het oplossen van complexe problemen. Volgens de constructivistische leertheorieën is het beeld van de lege vaten dus foutief. Dit blijkt bijvoorbeeld uit het feit dat leerlingen problemen kunnen oplossen op een andere dan de aangeleerde manier en uit het feit dat leerlingen aan wie de correcte leerinhoud onderwezen werd toch foutieve procedures hanteren.

b. Leren is cumulatief maar niet lineair

Het cumulatieve karakter van leren zit eigenlijk al vervat in het constructieve: leerlingen construeren nieuwe kennis op basis van wat ze al weten en kunnen. Voorkennis is een belangrijke factor bij het leren. Leerlingen bouwen steeds voort op en gebruiken kennis en vaardigheden waarover ze reeds beschikken (inclusief informele kennis). Toch mag men hieruit niet concluderen dat leren een mooi lineair proces is. Omdat leren niet lineair is en in verschillende richtingen kan verlopen, is het niet zo dat het leren van bepaalde concepten moet uitgesteld worden tot een bepaalde leeftijd of totdat bepaalde elementaire concepten beheerst worden. Mensen van alle leeftijden en ontwikkelingsniveaus gebruiken en verfijnen voortdurend hun concepten. Er is hierbij geenszins sprake van een vastliggend hiërarchisch parcours van het ene concept naar het volgende. Een belangrijke les hieruit is dat klassikale instructies (en zeker drill oefeningen) in principe

niet aansluiten bij bestaande schema's van alle leerlingen en dat er dus geen ideale manier van onderwijzen bestaat voor alle leerlingen.

c. Leren is zelfregulerend

Een andere invalshoek van het actieve en constructieve karakter van leren betreft het feit dat leren zelfgeregulerend is. Het gaat in deze over de activiteiten in verband met het beheeren en het bewaken van het leerproces door de leerling zelf, de zogenaamde metacognitieve vaardigheden (zie later in dit hoofdstuk). Leerlingen moeten leren om zelf de controle over hun eigen leren in handen te nemen.

d. Leren is doelgericht

Leren zal slechts succesvol zijn als de lerende zich bewust is van het zelfbepaalde doel waar hij of zij naar toe werkt. Het leren is met andere woorden doelgericht en zal het meest succesvol zijn als de leerlingen hun eigen doelen bepalen en nastreven. Momenteel is dat echter slechts beperkt het geval; de leer- en vormingsgebieden en de einddoelstellingen die nagestreefd moeten worden, zijn wettelijk vastgesteld. Dit houdt in dat de leerlingen vaak geconfronteerd worden met leerdoelen van buiten af. Het is dan erg belangrijk dat de leerlingen zich deze doelen eigen maken, dat ze zich ermee kunnen identificeren.

e. Leren is context- en situatiegebonden

Denk- en leerprocessen verlopen het best wanneer voortdurend interactie plaatsvindt met anderen en met de eigen culturele omgeving. Leren vindt immers niet in een vacuum plaats. Dit betekent bijvoorbeeld dat rekenopgaven niet geïsoleerd moeten worden tot rijtjes maar zoveel mogelijk geplaatst moeten worden in een levensechte situatie waarmee de kinderen in het dagelijks leven geconfronteerd worden. De lessen moeten dus een duidelijke band hebben met de realiteit. De kennis die erin aangeboden wordt, moet bruikbare, functionele kennis zijn. Dit geldt ook voor de illustratieve verduidelijking die er aan toegevoegd wordt.

f. Leren is collaboratief

De traditionele onderwijs- en evaluatiebenaderingen deden voornamelijk een beroep op individuele en onafhankelijke taken. Leerlingen moeten echter samen een 'lerende gemeenschap' vormen waar men verwacht wordt respectvol te luisteren naar elkaar, te reflecteren over en verder te bouwen op mekaars ideeën en kritisch te denken over uitgangspunten en argumenten van andere gezichtspunten. Een cruciaal aspect van het hedendaagse leren is de sociale interactie met anderen tijdens de leeractiviteiten. Samen met de leerkracht en met klasgenoten doorloopt de leerling het leerproces. Het werken in groep faciliteert het leren op verschillende manieren. Het gemeenschappelijk werken aan een bepaalde taak draagt bijvoorbeeld bij tot de kennisname van verschillende effectieve denkstrategieën, wederzijdse constructieve terugkoppeling, een waardering van het samenwerken met anderen en ondersteuning bij het verwerven van moeilijke en complexe vaardigheden en kennis.

Wanneer men de onderwijsbenadering baseert op sociaalconstructivistische leertheorieën zijn een aantal wijzigingen in het te onderwijzen curriculum noodzakelijk. In de volgende paragraaf gaan we dieper in op dit aspect van de hedendaagse onderwijsbenadering. Eerst staan we op basis van de aangehaalde kenmerken van leren stil bij een aantal kritieken op de huidige onderwijspraktijk.

Even stilstaan... 1.2: Kritieken op de huidige onderwijspraktijk

1. Zonder het hele onderwijssysteem over één kam te willen scheren; kan u akkoord gaan met volgende kritieken op de huidige onderwijspraktijk? Waarom wel/niet?
 - ✓ Informatieoverdracht is het dominante onderwijsleermodel.
 - ✓ Er wordt te weinig rekening gehouden met de voorkennis van leerlingen.
 - ✓ De onderwijsdoelen worden door de leerlingen niet als de hunne ervaren.
 - ✓ De invulling van de lessen is vaak wereldvreemd.
 - ✓ Leren wordt teveel gezien als een individuele activiteit.
2. Kan u op basis van de sociaal-constructivistische visie op leren nog andere kritieken op het huidige onderwijsgebeuren formuleren?

2. Het curriculum

Voorgaande elementen impliceren dat men van de lerende andere vaardigheden en eigenschappen verwacht dan voorheen. Het betreft naast een zekere basiskennis het onafhankelijk verwerken van informatie, oplossen van onvoorziene problemen, communiceren, samenwerken met anderen, etc. Het onderwijs moet leerlingen opleiden tot denkende, autonome, zich aanpassende en zelfregulerende personen. Van een dergelijke persoon verwachten specifieke vaardigheden en kenmerken. De hedendaagse onderwijsbenaderingen nemen afstand van het uitgangspunt dat intelligentie en capaciteiten ééndimensioneel en vastliggend zijn. Naast de traditioneel benadrukte capaciteiten (verbaal-linguïstisch en logisch-mathematisch) worden in de hedendaagse onderwijsbenaderingen dan ook andere belangrijke curriculuminhouden benadrukt: visueel-spatieel, het bewegingsgevoel, muzikaal, intrapersoonlijk en interpersoonlijk. Te denken valt aan de volgende vier groepen vaardigheden die telkens samengaan met een bepaalde rol en die telkens als een afzonderlijk curriculumdomein kunnen beschreven worden: het cognitieve domein, het metacognitieve domein, het affectieve domein en het sociale domein¹. Belangrijk hierbij is dat er grote verschillen bestaan tussen leerlingen in de wijze waarop en de snelheid waarmee deze verschillende domeinen ontwikkeld worden.

a. Het cognitieve domein

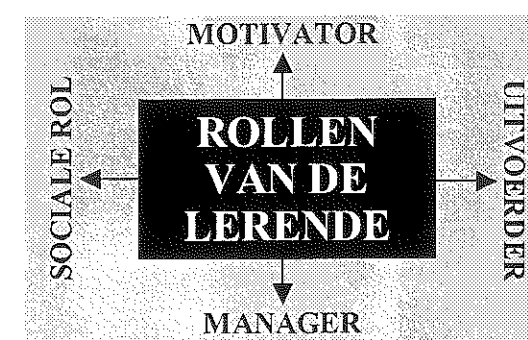
Binnen dit domein gaat het om cognitieve competenties zoals een probleem oplossen, kritisch denken, vragen formuleren, geïnformeerde oordelen maken, observeren en onderzoeken, uitvinden en creëren... Leerlingen voeren denkactiviteiten uit die voornamelijk gericht

¹ Gebaseerd op: Birenbaum, 1995; Clemens et al., 1999; Herman et al, 1992;; Shepard, 1996; Rolheiser, et al., 2000; Janssens et al., 2000.

zijn op het begrijpen en integreren van nieuwe kennis. Hierbij past de rol van daitvoerder: degene die bijvoorbeeld rekent, leest, herhaalt, samenvat, concludeert, probeert om inzicht te krijgen.

b. Het metacognitieve domein

Hierbij gaat het om het sturen en monitoren van het leerproces. In deze rol worden als het ware de managementfuncties vervuld. Te denken valt aan metacognitieve competenties zoals het stellen van leerdoelen, het maken van een plan, het verifiëren van eigen oplossingen, zelfreflectie en zelfevaluatie (het controleren of de einddoelen bereikt zijn).



Figuur 1.2: Rollen van de lerende

c. Het affectieve domein

Binnen het affectieve domein valt te denken aan elementen als doorzettingsvermogen, verantwoordelijkheid, initiatief nemen, doelmatigheidsbeleving, zelfstandigheid, flexibiliteit en het hoofd bieden aan frustrerende situaties. Het gaat om de affectie bij het leren. Hierbij past de rol van (zelf)motivator: 'kan ik het wel (in de voorziene tijd)', 'wil ik het wel', 'wat is er leuk aan'? Leren wordt sterk bepaald door dergelijke affectieve factoren.

d. Het sociale domein

Wanneer we de aandacht vestigen op het feit dat er steeds meer met anderen samen wordt geleerd, kunnen we nog een vierde domein benadrukken: het sociale. Sociale competenties maken immers een belangrijk deel uit van het hedendaagse curriculum en kunnen dus een doel op zich zijn. Te denken valt aan discussies en conversaties leiden, overtuigen, luisteren, samenwerken en in groep functioneren. Hierin schuilt de sociale rol van de lerende.

3. Het onderwijsleerproces

Om een dergelijk curriculum succesvol te onderwijzen, moeten een aantal belangrijke richtlijnen in rekening gebracht worden. De hedendaagse onderwijsbenaderingen kenmerken zich dan ook door specifieke verwachtingen ten aanzien van het eigenlijke onderwijsleerproces. Twee herhaaldelijk benadrukte elementen zijn het creëren van krachtige leeromgevingen en de promotie van diepgaand leren. Van beide geven we kort de essentie weer.

a. Creëren van krachtige leeromgevingen

Zoals we hoger beschreven, hebben de hedendaagse onderwijsdoelstellingen sterke repercussies voor het onderwijscurriculum en het onderwijsleerproces. Een mogelijkheid om hieraan tegemoet te komen, bestaat erin zich te richten op het realiseren van een 'krachtige leeromgeving' (Verschaffel & Vermunt, 1998; Verschaffel, 1995; Dochy & Segers, 1999; Janssens et al., 2000; Herman et al., 1992; De Corte, 1990). Bij het creëren van krachtige leeromgevingen moet bijzondere aandacht worden besteed aan de onderwijsinhoud, de onderwijsmethoden, de sequentie van de leertaken en de sociale context van het leren. Volgende elementen kenmerken de invulling van deze vier dimensies van een krachtige leeromgeving.

- ✓ **Samenwerking.** Studenten worden uitgedaagd om samen met elkaar actief te leren. De samenwerking bestaat concreet bijvoorbeeld uit het in groep oplossen van problemen en uit het voeren van klassikale gesprekken waarin verschillende zienswijzen en werkwijzen aan bod komen. In een krachtige leeromgeving zijn tevens modellen en voorbeelden van expertgedrag (het 'goede' voorbeeld) aanwezig in de rol van de leerkracht en van de medeleerlingen.
- ✓ **Toepassingsperspectief.** Er is een duidelijk zicht op de functionaliteit van het leren en het geleerde. Eerder dan eenvoudige en afzonderlijke vaardigheden te benadrukken, moet men leerlingen tijdens het onderwijsleerproces onderdompelen in uitdagende en toepassingsgerichte contexten die complex en holistisch denken vereisen. Belangrijk is deze holistische taken in toenemende moeilijkheidsgraad aan te bieden zodat leerlingen gemotiveerd en aangemoedigd worden en komen tot een sterke doelmatigheidsbeleving en tot een realistisch zelfvertrouwen. Het is hiertoe van essentieel belang de ervaringen en voorkennis van leerlingen als uitgangspunt te nemen en de leerlingen eerst een brede oriëntatie te geven op het leerdoel alvorens de aandacht te richten op deelvaardigheden. Door het leggen van verbanden met de nieuwe leerstof kan de voorkennis verfijnd worden. Het gaat dus om 'knowlegde building' in plaats van 'knowlegde reproduction'.
- ✓ **Concrete, realistische context.** Het leren is gesitueerd in een realistische context met een zo hoog mogelijke authenticiteit. Het is dan ook aangewezen gebruik te maken van realistische, authentieke taken die samenwerking en vakoverschrijdende kennis activeren. Krachtige leeromgevingen zijn onderwijssettings waarbij wordt uitgegaan van authentieke situaties, problemen, casussen en opdrachten die tot doel hebben de leerlingen kennis en vaardigheden te leren toepassen.
- ✓ **Intrinsieke leermotivatie** Het leren is gemotiveerd vanuit een eigen belangstelling van de leerling en niet vanwege de externe beloning waartoe succesvol leren zou kunnen leiden. Meer algemeen gesteld komt dit principe erop neer dat voldoende aandacht wordt gegeven aan dynamisch-affectieve aspecten van het onderwijsleerproces.
- ✓ **Zelfstandig leren.** Studenten worden geacht zelfstandig hun leren te kunnen sturen en controleren binnen de leeromgeving. Er is niet noodzakelijk een objectieve

waarheid met betrekking tot de leerstof, maar leerlingen worden geacht hun eigen perspectief te ontwikkelen en de vooropgestelde doelen te bereiken.

- ✓ **Begeleiding.** De rol van de docent zal van coachende en begeleidende aard zijn in plaats van het overdragen van informatie. Het feit dat leren omschreven werd als een actief en constructief proces impliceert dus niet dat dit leerproces best zonder enige sturing van buitenaf plaatsvindt. Een aantal vormen van begeleiding die hun waarde reeds bewezen hebben, zijn modelleren, scaffolding, coaching, articulatie en reflectie.

b. Promotie van diepgaand leren

Zoals gesteld is samen met de onderwijsdoelen ook de visie op leren geëvolueerd. Een andere benadering van het onderwijscurriculum brengt de manier waarop leerlingen leerstof (informatie) verwerken en de leerstrategieën die zij hierbij gebruiken samen in het concept 'cognitief verwerkingsniveau' (Struyf et al., 1999; Herman et al., 1992). Vaak wordt in dit opzicht een onderscheid gemaakt tussen een oppervlakkige en een diepgaande verwerking van leerstof. Bij diepgaande verwerking wil men de betekenis van de informatie achterhalen en maakt men gebruik van een veelheid aan cognitieve strategieën om zich de informatie eigen te maken (zoals discrimineren van hoofd- en bijzaken, relateren, concretiseren, structureren, en dergelijke). Bovendien hanteert men metacognitieve strategieën. Het diepgaande verwerkingsniveau vereist dus een aantal vaardigheden en kenmerken van de leerder zoals hoger beschreven.

Diepgaande verwerking wordt afgezet tegen een oppervlakkige verwerking die eerder kenmerkend is voor de traditionele onderwijsbenadering. De tegenstelling tussen diepgaande versus oppervlakkige verwerking wordt dus gekenmerkt door het al dan niet toepassen van denkactiviteiten als hoofd- en bijzaken onderscheiden, het kritisch volgen van een argumentatie en het zoeken van verbanden. De oppervlakkige verwerking wordt gekenmerkt door de aandacht die gegeven wordt aan het leren van de tekst en aan het reproductief concept van studeren (van buiten leren). De leerstof wordt beschouwd als een aaneenschakeling van losse feiten waarvan men zoveel mogelijk details wil onthouden.

Uit deze twee aandachtspunten omtrent het succesvol inrichten van onderwijsleerprocessen blijkt dat onderwijs niet meer wordt beschouwd als het overdragen van kennis, maar als het bevorderen van leerstrategieën bij leerlingen. De leerkracht krijgt de rol van mentor of coach die de lerende kansen biedt om zijn voorkennis te gebruiken en om nieuwe elementen te begrijpen.

4. Evaluatie

Ook het evalueren moet aan bovenstaande onderwijsvisie worden aangepast. In de volgend paragraaf beschrijven we uitvoerig de veranderingen in de evaluatiebenaderingen en -methoden die de overgang van de traditionele naar de hedendaagse onderwijsbenadering kenmerken. We stellen ons tevens de vraag op welke manier de evaluatie van leerlingen moet worden ingevuld

om bij de voorgestelde onderwijsdoelen, het curriculum en het onderwijsleerproces aan te sluiten.

2.2 Evoluerende evaluatiebenaderingen: van toetspraktijk naar een veelzijdige, geïntegreerde assessmentcultuur

De verschillen tussen de opvattingen over leren en de onderwijsdoelen in de twee beschreven periodes worden gereflecteerd in de geïmplementeerde evaluatiemethoden. Een overschakeling van de ene naar de andere onderwijsbenadering met een status-quo wat de evaluatie betreft, zou immers onherroepelijk leiden tot verwarring bij leerlingen en tot het verdwijnen van de meerwaarde van de innovatie. Men moet steeds in het achterhoofd houden dat de evaluatie (momenteel dus vaak toetsen en examens) het sterkst sturende aspect is van de leeractiviteiten van leerlingen (Dochy & Segers, 1999). Hieruit kunnen we concluderen dat de evaluatiebenadering mee moet evolueren indien men inhoudelijk en didactisch tot nieuwe benaderingen en doelstellingen wenst te komen.

Kortom: ook op het vlak van evaluatie hebben zich verschillende ontwikkelingen voorgedaan. Hieronder spreken we van een traditionele en een alternatieve evaluatiebenadering. Op die manier wordt duidelijk waarvoor zogenaamde 'alternatieve' evaluatievormen effectief een alternatief vormen. We zullen voornamelijk aandacht hebben voor de *evaluatiecultuur* waarin de alternatieve evaluatiemethoden gesitueerd kunnen worden. Deze vormt immers een ruimer kader waarin het verdere verloop van het beschrijven van alternatieve evaluatievormen in dit handboek vorm zal krijgen. Na de beschrijving van de kenmerken van de traditionele en de alternatieve evaluatiebenaderingen vatten we de belangrijkste evoluties samen.

Eerst nog even dit. De benaming *traditionele toetscultuur* impliceert niet dat de bedoelde evaluatiebenadering reeds tot het verleden behoort. Al te vaak wordt de huidige onderwijspraktijk immers nog gekenmerkt door de zogenaamde toetscultuur. Gestandaardiseerde toetsen zullen bovendien niet zomaar van het onderwijstoneel verdwijnen (Fisher & King, 1995). Verschillende mensen geloven nog in de meerwaarde van traditionele toetsentout court. Bovendien zijn de traditionele toetsen wel degelijk geschikt om de realisatie van bepaalde onderwijsdoelstellingen te evalueren. Met deze nuancering in het achterhoofd beschrijven we in de volgende paragrafen de voornaamste kenmerken van de traditionele toetscultuur.

2.2.1 Van een traditionele toetscultuur...

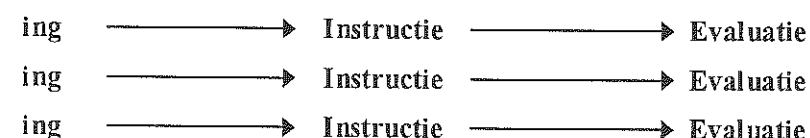
We richten ons in het onderstaande eerder algemeen op de evaluatiecultuur die het gebruik van traditionele toetsen kenmerkt en minder op de eigenlijke traditionele evaluatie-instrumenten. Indien relevant geven we eveneens aan welke de belangrijkste kritieken zijn op deze evaluatiebenadering.

1. Evaluatie als afsluiting van een leerstofdeel

Een belangrijk vertrekpunt is de vaststelling dat in de (traditionele) evaluatieliteratuur drie fasen worden onderscheiden in het onderwijsleerproces (Struyf et al., 1999):

- de planning van de instructie;
- het geven van de instructie;
- nagaan wat de leerlingen geleerd hebben, de evaluatie.

De leerkracht plant, geeft een instructie en in een daarvan los gekoppelde evaluatie wordt daarna nagegaan wat de leerlingen geleerd hebben. Deze fasen volgen elkaar in herhaalde cycli op zodat de evaluatie telkens gescheiden is van het leer- en onderwijsproces. De evaluatiefase komt dus telkens aan het einde van een leerperiode en het studeren staat in het teken van de evaluatie.



Figuur 1.3: De traditionele relatie tussen instructie en evaluatie

2. De traditionele evaluatie-instrumenten

De specifieke evaluatie-instrumenten die gebruikt worden, zijn zeer typerend voor de traditionele onderwijsbenadering. De traditionele tests bestaan meestal uit een vraag-antwoordformat zoals multiple-choice, waar-vals vragen, juist-foutvragen en korte-antwoord-vragen.

Figuur 1.4: Voorbeelden van multiple-choice- en waar-vals vragen

Multiple-choicevragen:

Het kookpunt van zuiver water, op zeeniveau en in een open bokaal is:

- 50 °
- 85°
- 100°
- 200°

Waar-vals vragen:

- ✓ Het Nederlandstalige begrip 'evaluatie' is ruimer dan het Engelstalige 'assessment' (waar/vals).
- ✓ Leren is cumulatief. Dit wil zeggen dat elke leerling een vastliggende hiërarchie van steeds complexere mentale schema's doorloopt (goed/fout).
- ✓ Het sociaal-constructivistische gedachtegoed is gebaseerd op cognitieve leertheorieën (correct/incorrect).

In het Vlaamse onderwijs is het gebruik van deze evaluatie-instrumenten niet zo ingeburgerd als in buitenlandse onderwijssystemen met hun gestandaardiseerde en centrale examens (zoals in Verenigde Staten, Engeland en Nederland). Wanneer we ons richten op de kenmerken van traditionele evaluatievormen moeten we concluderen dat niet zozeer de typische evaluatie-instrumenten maar wel de achterliggende evaluatiecultuur in Vlaanderen teruggevonden wordt.

Over die cultuur kunnen we het volgende stellen. Vaak moet de leerling het juiste antwoord kiezen uit een aantal opgegeven alternatieven of het korte antwoord zelf construeren. Een basisregel in de toetscultuur is dat leerlingen individueel en zonder hulpmiddelen (zoals een rekenmachine of woordenboek) getoetst worden (Struyf et al., 1999). De evaluatie gebeurt bovendien vaak onder tijdsdruk. De traditionele evaluatiebenadering richt zich voornamelijk op het toetsen van kennis en fundamentele vaardigheden die op basis van herhaalde drill oefeningen aangeleerd worden. Struyf (1999) en Fisher & King (1995) schrijven dat het meest geëvalueerde domein het cognitieve is en dat binnen dat domein vaak nog wordt toegespitst op het niveau van de kennis. De focus ligt dus op de evaluatie van de basiskennmerken en eenvoudige vaardigheden terwijl complexe denk- en probleemoplossingsvaardigheden verwaarloosd worden (Shepard, 1996).

Net deze inhoudelijke beperktheid en de slechte overeenstemming tussen de inhoud van de toets enerzijds en het curriculum en de instructie anderzijds maken dat deze evaluatievormen veel kritiek oogsten. De vraag is dan ook of dergelijke gestandaardiseerde, meerkeuze-toetsen betekenisvolle evaluatiegegevens genereren. Critici beweren alvast van niet (Shepard, 1996; Herman et al., 1992; Arter & Bond, 1996). Voornamelijk de kritiek dat de nadruk te zeer ligt op basisvaardigheden en dat leerlingen eenzijdig worden voorbereid op het succesvol afleggen van gestandaardiseerde toetsen (teaching to the test) wordt benadrukt. Dit geldt des te meer naarmate het belang van de evaluatieresultaten groter wordt (zoals bij de vergelijking van de 'kwaliteit' van individuele scholen). De traditionele toetsvormen oogsten evenwel nog andere kritieken. Studenten uit sociaal lagere klassen zouden bij deze toetsen benadeeld worden omdat de toetsen geen beroep doen op de specifieke context waarin die leerlingen leven. Bovendien doen de traditionele multiple-choice toetsen te weinig een beroep op complexe cognitieve en meta-cognitieve vaardigheden. Concluderend: de traditionele evaluatie-instrumenten slagen er niet in significante leeruitkomsten te evalueren en brengen daardoor het vooropgestelde curriculum en de geplande instructies in gevaar (Herman & Knuth, 1991).

3. De functie en vormgeving van evaluatie

De functie van het evalueren bestaat voornamelijk uit het beoordelen van leerlingen. Deze beoordeling vormt de basis om leerlingen te oriënteren, te selecteren. De beoordeling is vaak normgerefeerd; men vergelijkt de resultaten van een leerling met de resultaten van andere leerlingen. Deze resultaten worden meestal kwantitatief gerapporteerd in de vorm van een cijfer. Ook de bespreking van de resultaten gebeurt aan de hand van cijfers. De beoordelende functie van evaluatie in de toetscultuur houdt tevens in dat de leerkracht verantwoordelijk is voor de

evaluatie. Deze stelt de toets op, kiest de items, controleert de afname van de toets en interpreteert en beoordeelt de antwoorden. De leerkracht heeft m.a.w. de volledige controle over het evaluatieproces. De leerlingen hebben daarentegen nauwelijks een inbreng in het evaluatiegebeuren. Zij bepalen slechts zelden wanneer de toets wordt afgenomen en wat de inhoud ervan zal zijn. Bovendien begrijpen ze vaak niet volledig wat de toetsscore betekent of hoe die tot stand is gekomen (Struyf et al., 1999; Janssens et al., 2000). Dit betekent ook dat de kans dat leerlingen iets leren uit de evaluatie gering is.

Met deze kenmerken van de traditionele evaluatiecultuur in het achterhoofd nemen we de tijd om even stil te staan bij de eigen onderwijscontext.

Even stilstaan... 1.3: Sporen van een traditionele evaluatiebenadering

Enkele vragen om de eigen evaluatiepraktijk eens kritisch te bekijken.

- ✓ Vindt de evaluatie steeds plaats na het afronden van een leerstofgeheel of is de evaluatie geïntegreerd in de instructie en het leerproces zelf?
- ✓ Bij het bespreken van het voor de huidige onderwijsbenadering kenmerkende curriculum maakten we een onderscheid tussen vier inhoudelijke domeinen. Komt elk van deze domeinen aan bod in uw evaluaties? Welk van deze domeinen krijgt het meeste aandacht? Welk het minste?

2.2.2 ...naar een assessmentcultuur

We stelden reeds dat er ten aanzien van evaluatie verschillende ontwikkelingen hebben plaats gevonden. Er wordt in de literatuur en in de praktijk gezocht naar andere vormen van evaluatie waarin bijvoorbeeld niet alleen de leerkracht maar ook de leerlingen nieuwe rollen vervullen. Daarnaast wordt gezocht naar manieren om deze evaluatievormen beter te integreren in het onderwijsleerproces. Hierna geven we een uitvoerige beschrijving van de evaluatieopvatting waartoe deze ontwikkelingen als het ware leid(d)en. Vaak verwijst men in dit opzicht naar de zogenaamde assessmentcultuur (Birenbaum, 1995; Struyf et al., 1999). Zoals uit figuur 1.5 blijkt, zijn verschillende belangrijke evoluties kenmerkend voor de overgang van een toetscultuur naar een assessmentcultuur.

In de volgende paragrafen richten we ons op een aantal specifieke aspecten van de assessmentcultuur: de functie en de doelen, de inhoud, de evaluatiemethoden, de verdeling van de verantwoordelijkheden, de integratie in de instructie en het rapporteren van resultaten.

Figuur 1.5: Van een toetscultuur naar een assessmentcultuur

Van een... toetscultuur		naar een... assessmentcultuur	
van een...		naar een...	
behavioristische visie op leren	↔	constructivistische visie op leren	
beoordelende functie	↔	diagnostiserende en begeleidende functie	
focus op reproduceren van kennis	↔	focus op construeren van kennis	
productevaluatie	↔	proces en productevaluatie	
schriftelijke toets	↔	veelzijdige evaluatie	
passieve rol van de leerlingen	↔	actieve rol van de leerlingen	
verantwoordelijkheid van de leerkracht	↔	gedeelde verantwoordelijkheden	
normgerichte beoordeling	↔	criteriumgerichte beoordeling	
kwantitatieve rapportering	↔	kwantitatieve en kwalitatieve rapportering	
cyclische en geïsoleerde instructie en evaluatie	↔	geïntegreerde instructie en evaluatie	

1. De functie/doelen van de evaluatie

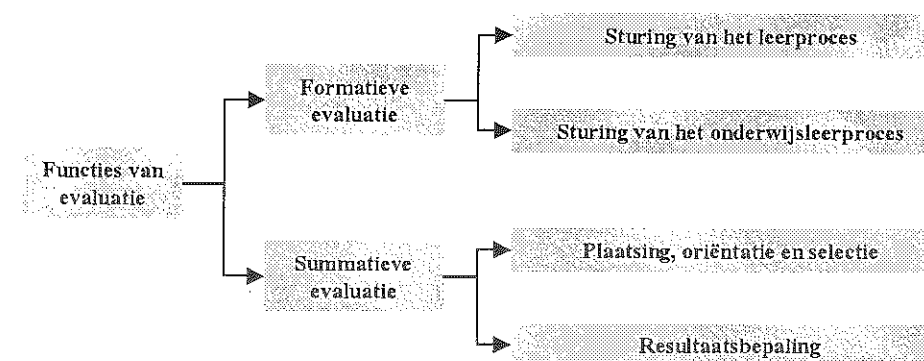
In vergelijking met de traditionele toetscultuur is de functie van evalueren geëvolueerd van beoordelen richting begeleiden en diagnosticeren. Er is dus sprake van een verschuiving van summatief naar formatief evalueren. Formatieve evaluatie omschreven we hoger als een tussentijdse vorm van evaluatie die de basis vormt voor de optimalisering van het onderwijsleerproces, d.i. van het leerproces van de leerlingen/studenten enerzijds en het onderwijzen van de leerkracht/docent anderzijds. Te denken valt in het eerste geval aan het opsporen van sterke en zwakke punten in de prestaties van de leerlingen. Op basis van dergelijke evaluatiegegevens kan men

indien nodig de remediëring en de bijkomende begeleiding van de leerling plannen. Er wordt met andere woorden gereflecteerd over de planning en de uitvoering van de instructie die het meest is aangewezen voor de betreffende leerling. Daarnaast gebruikt de leerkracht de evaluatiegegevens om het onderwijsleerproces bij te sturen en om wijzigingen aan te brengen in het didactisch handelen. De evaluatiefuncties binnen een assessmentcultuur reiken dus duidelijk verder dan het beoordelen van leerlingen met het oog op selectie en resultaatsbepaling in de traditionele evaluatiecultuur (zoals overgaan naar een volgend schooljaar en toekennen van diploma's). Wanneer men zich toch richt op de eindprestaties van de leerlingen, worden tevens de gemaakte vooruitgang en de getoonde inzet geëvalueerd.

Een aantal functies van evaluatie zijn zowel kenmerkend voor traditionele als voor alternatieve evaluatievormen. Door te evalueren worden standaarden vooropgesteld en gecommuniceerd naar de betrokkenen, omdat wordt aangegeven wat men verwacht van een goede prestatie. De inhoud van de evaluatie leert leerlingen wat belangrijk is en wat al dan niet bijkomende aandacht verdient. Bovendien mag ook het motiverende effect van het evalueren op leerlingen niet uit het oog worden verloren. Hierbij toch even dit: hoewel deze functies gemeenschappelijk blijken, zijn de vorm en de inhoud van een evaluatie binnen de assessmentcultuur van die aard dat duidelijk andere standaarden, prioriteiten en verwachtingen gecommuniceerd worden dan binnen de traditionele evaluatiecultuur.

Binnen een assessmentcultuur krijgen formatieve evaluatiefuncties de bovenhand op summatieve functies. Om de bespreking van de evaluatiefuncties af te ronden presenteren we een samenvattend schema (zie fig. 1.6) met een aanvullende denkoefening.

Figuur 1.6: De evaluatiefuncties binnen een assessmentcultuur



Even stilstaan... 1.4: Evaluatiefuncties in de huidige onderwijscontext

1. Werp even een blik op de voor jou meest vertrouwde evaluatiecontext. Beantwoord vervolgens volgende vragen over de functies van evaluatie in die situatie.

- ✓ Wat is de verhouding tussen de formatieve en summatieve evaluatie. Welke van beide komt het meest aan bod? Wat concludeert u op basis van de verhouding?
- ✓ Welke van de vier functies van evaluatie vind u zelf het meest belangrijk? Waarom?

2. In het jaarverslag van de onderwijsinspectie (2001, p. 4) lezen we het volgende: Een aangepaste evaluatie en rapportering blijft de zwakke schakel in het differentiatieproces en hypothekeert het streven naar adaptief onderwijs. Scholen missen kansen om de vorderingen bij leerlingen vast te stellen en om op basis van de evaluatiegegevens verdere gedifferentieerde leertrajecten uit te tekenen. De evaluatiegegevens worden nog te weinig geanalyseerd en geïnterpreteerd. Ze blijven te veel doel op zich en in veel mindere mate de basis om te kunnen differentiëren in doelen en leerinhouden.

- ✓ Wat leert ons dit over de huidige functies van evaluatie?
- ✓ Bent u het eens met de stelling van de inspectie?

2. De inhoud van de evaluatie

In vergelijking met de toetscultuur is de inhoud van de evaluatie in de assessmentcultuur uitgebreider. In plaats van louter gericht te zijn op eenvoudige basiskennis en -vaardigheden, worden eerder complexe denkprocessen, vaardigheden en attitudes bij leerlingen/studenten geëvalueerd. Het construeren van kennis is belangrijk, niet het reproduceren ervan. Daarnaast komen ook het sociaal-emotionele en het psychomotorische domein in de evaluatie aan bod (cf. het affectieve en sociale domein). De inhoud van evaluaties binnen een assessmentcultuur omvat dus elk van de vier domeinen die we in wat voorafgaat onderscheidden in het curriculum van de hedendaags onderwijsbenadering: het cognitieve, het metacognitieve, het affectieve en het sociale.

Onderwijsdoelen zoals competenties en praktische vaardigheden zijn niet te evalueren aan de hand van traditionele 'paper-and-pencil test', maar vragen om een integrale complete toetsing, die indien mogelijk in een realistische context geplaatst wordt. De kanteling naar een assessmentcultuur kan dus aangezien worden als een reactie op de inhoudelijke tekortkomingen van de traditionele evaluatiemethodes om de veranderde onderwijsdoelen te evalueren (Struyf et al., 1999; Arter & Bond, 1996).

3. De evaluatiemethoden

Met het oog op de eigenheid van de doelstellingen die men binnen een assessmentcultuur wil evalueren, zijn de traditionele toetsen niet geschikt. Binnen een assessmentcultuur worden dan ook verschillende alternatieve methoden gebruikt om evaluatiegegevens over leerlingen te verzamelen (Struyf et al., 1999). Te denken valt aan vragenlijsten, (groeps)gesprekken, verslagen van opdrachten, reflecties, en dergelijke. Doorgaans kunnen we stellen dat de evaluatiemethoden evolueren van het stellen van vragen met korte (vastgelegde) antwoorden richting het geven van opdrachten en observeren. De onderwijsgevende kan het geobserveerde gedrag uitvoerig beschrijven of hij kan bij de observatie gebruik maken van een checklist of een beoordelingschaal. Vaak wordt er bovendien gezocht naar evaluatievormen die gebruik maken van zinvolle, authentieke en uitdagende opdrachten. Dit betekent bijvoorbeeld ook dat leerlingen vaak verschillende hulpmiddelen (die eveneens in het dagelijkse leven gebruikt worden) mogen hanteren en dat het mogelijk is dat leerlingen onderling en met de leerkracht samenwerken. We stelden dat een belangrijk doel binnen de assessmentcultuur bestaat uit het begeleiden van leerlingen. Dit uit zich in het feit dat men tracht de evaluatie zoveel mogelijk te laten aansluiten bij de indi-

viduele verschillen tussen leerlingen (capaciteiten, leerstijlen, talenten, interesses, persoonlijkheid).

De assessmentcultuur kent dus een verscheidenheid aan evaluatievormen. Zo kan men telkens die evaluatievorm kiezen die het best bij de vooropgestelde onderwijsdoelstellingen en kenmerken van de leerling aansluit. In deel 2 bespreken we de kenmerken van alternatieve evaluatievormen en in deel 4 beschrijven en illustreren we de verschillende alternatieve evaluatiemethoden.

4. Delen van verantwoordelijkheid

In tegenstelling tot de toetscultuur is binnen de assessmentcultuur wel een belangrijke en actieve rol weggelegd voor de leerlingen. De verantwoordelijkheden ten aanzien van de evaluatie worden gedeeld door de leerkracht en de leerlingen. Struyf et al. (1999, p. 8) beschrijven deze verantwoordelijkheden als volgt: "De leerkracht moet verder leren mogelijk maken en bevorderen; hij moet de leerlingen leren om zelf verantwoordelijkheid te dragen voor hun eigen leerproces. Op die manier ontwikkelt de leerkracht bij de leerlingen een grotere mate van zelfstandigheid en vertrouwen, die de basis vormen van 'levenslang leren'. De leerling is een actieve participant die mee verantwoordelijk is voor het evaluatieproces. Leerlingen worden aangemoedigd om over hun werk en leerproces na te denken."

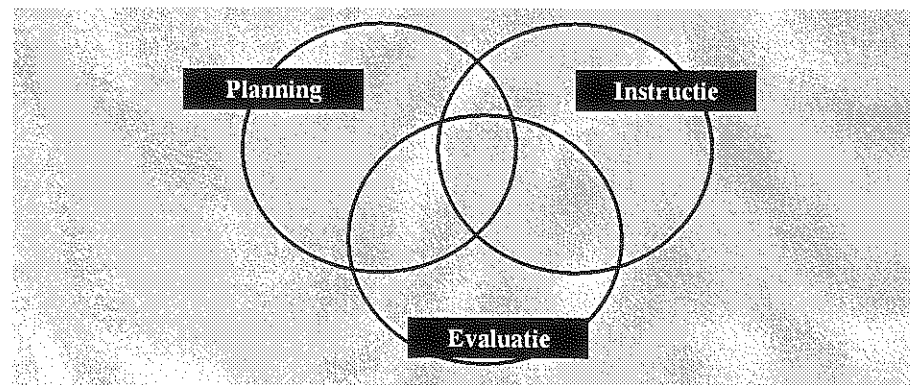
Kenmerkend voor de assessmentcultuur is dat de inbreng van de lerende bij de evaluatie vergroot wordt. Leerlingen worden opgevat als zelfstandige, autonome en initiatiefrijke individuen die in grote mate zelf hun leerproces sturen (Dochy & Segers, 1998). Op die manier komen evaluaties niet meer volledig in handen van de leerkracht te liggen, maar worden de leerlingen zelf verantwoordelijk voor de evaluatie van hun eigen leerproces. Leerlingen moeten bijvoorbeeld in staat zijn hun eigen sterke en zwakke punten op te sporen om in functie hiervan hun verdere leerproces te plannen. Dit vergt het verwerven van verscheidene vaardigheden zoals het vooropstellen of tenminste (h)erkennen van (leer)doelen, het ontwikkelen van evaluatiecriteria en standaarden, het reflecteren en feedback geven. Het dragen van verantwoordelijkheid maakt dat leerlingen meer betrokken worden bij de leerstof en gemakkelijker gebruik gaan maken van diepgaande verwerkingsstrategieën (Topping, 1998; Longhurst & Norton, 1997). Uit onderzoek is bovendien gebleken dat leerlingen waarderen dat ze actief bij het leer- en evaluatieproces worden betrokken (Struyf et al., 1999). De verantwoordelijkheid die leerlingen dragen, is tevens indicatief voor de integratie tussen instructie en evaluatie die in de volgende paragraaf zal besproken worden. Die integratie is interessant omdat een dergelijke verantwoordelijkheid van de leerlingen ook op zich bepaalde onderwijsdoelen kan dienen; ze is met andere woorden niet alleen een doel op zich maar ook een middel tot.

5. Het integreren van instructie en evaluatie

Ook in de assessmentcultuur kunnen de drie fases onderscheiden worden die de toetscultuur kenmerken. Verschillend van de traditionele evaluatiecultuur is echter dat de verschillende fases

meer geïntegreerd verlopen. De grenzen tussen instructie en evaluatie vervagen en de evaluatie wordt meer geïntegreerd in het onderwijsleerproces zelf. Deze evolutie sluit aan bij het vooropstellen van het begeleiden van leerlingen en het optimaliseren van de instructie als belangrijke evaluatiedoelen.

Figuur 1.7: De integratie van instructie, planning en evaluatie



Bij het bespreken van de integratie van evaluatie met onderwijzen en leren als kenmerk van alternatieve evaluatievormen besteden we aandacht aan de motieven voor deze integratie (zie deel 2). Daarnaast komen enkele richtlijnen bij het realiseren ervan aan bod.

6. Het rapporteren van resultaten

De assessmentcultuur kenmerkt zich door een meer kwalitatieve rapportering over de leerlingresultaten. In plaats van een kwantitatief cijfer (zoals 7 op 10) maakt men gebruik van kwalitatieve beoordelingsschalen en leerlingprofielen (Struyf et al., 1999). Die geven namelijk een nauwkeurige analyse van de sterke en zwakke punten van de leerlingen en beschrijven waarom een leerling een bepaalde score kreeg. Indien ze al gebruikt worden, worden puntenrapporten in een assessmentcultuur dan ook aangevuld met andere gegevens. Te denken valt aan een woordelijke uitleg, een verbaal rapport, portfolio's en dergelijke. Belangrijk in de assessmentcultuur is verder de criteriumgerichte werkwijze bij het rapporteren van de beoordelingen van de leerlingenprestaties. Dit betekent dat de prestatie van een leerling wordt afgewogen tegen een vooropgesteld criterium (bij wijze van voorbeeld: 80% van de opgaven correct oplossen). Deze criteriumgerichte beoordeling kan worden afgezet tegen de normgerichte beoordeling waar men de prestaties van een leerling bekijkt in relatie tot de prestaties van andere leerlingen of een normgroep. Ook de feedback die leerlingen na een evaluatie krijgen, benadrukt vooral de beheersing

en de vooruitgang van de leerling zelf en niet zozeer de vergelijking met klasgenoten. Op die manier kan men vermijden dat vormen van competitie en sociale vergelijking een negatief effect hebben op de intrinsieke motivatie van de leerlingen.

Bij het bespreken van het evaluatieproces (c.q. van de communicatie van evaluatiegegevens) in deel 3 komen de verschillende aspecten van het rapporteren van evaluatiegegevens binnen een assessmentcultuur verder aan bod.

2.2.3 Samenvattend: de belangrijkste evoluties nader bekeken

In de huidige onderwijspraktijk vindt men vaak nog toetsen die gericht zijn op de reproductie van kennis. Zoals uit bovenstaande kenmerken kan blijken, gaat een assessmentcultuur echter verder dan dat. De evolutie van de functies en de inhoud van evaluatie stimuleert tot het initiëren van een evaluatiecultuur met aangepaste evaluatiemethoden, met veranderde rollen voor leerkrachten en leerlingen, met een geïntegreerde evaluatie en instructie en met een kwalitatieve rapportering van evaluatieresultaten. De zogenaamde assessmentcultuur die deze evoluties bundelt, kunnen we samenvattend als volgt beschrijven:

In de assessmentcultuur wordt de integratie tussen instructie en assessment sterk benadrukt. De leerling is een actieve participant die mede verantwoordelijkheid draagt voor het leerproces, die zichzelf evalueert, die reflecteert en samenwerkt in een aanhoudende dialoog met de leerkracht. De evaluatie zelf kan verschillende vormen aannemen. De taken zijn interessant, betekenisvol, authentiek, uitdagend en vergen telkens een vorm van onderzoek. De leerlingen participeren in de ontwikkeling van criteria en standaarden die zullen gehanteerd worden om hun prestaties te beoordelen. Zowel het proces als het product worden geëvalueerd en gerapporteerd op een manier die kan variëren van een allesomvattend cijfer tot het opstellen van een profiel of uitvoerige beschrijving.

In de evolutie van traditionele evaluatiebenaderingen naar de assessmentcultuur, kunnen we een drietal (deels overlappende) veranderingen onderkennen, namelijk (1) van cognitieve naar multidimensionele assessment, (2) van toetsen naar multiple assessment en (3) van geïsoleerde naar geïntegreerde assessment. We bespreken kort deze belangrijke evoluties.

1. Van cognitieve naar multidimensionele assessment

Traditioneel lag een eng concept van de menselijke intelligentie aan de basis van de visies op leren. Kulieke et al. (1990) schrijven dat het traditionele onderwijs eenzijdig verbaallinguïstische en logisch-mathematische vaardigheden benadrukte. Hierdoor werd aan leerlingen impliciet de boodschap gegeven dat dit de enige belangrijke dimensies waren. Er zijn echter verschillende andere kennisvormen, talenten en vaardigheden die ons leven verrijken en ons helpen effectief te reageren op onze omgeving. Het incorporeren van deze elementen in het onderwijs vereist dat deze elementen ook in het evaluatieproces worden opgenomen. Het initiëren van een concept van 'multidimensionele assessment' betekent dat de evaluatie van leerlingen gebaseerd

is op een breder concept van intelligentie, vaardigheden en leren. Zo'n concept heeft niet alleen aandacht voor linguïstische en logistische aspecten, maar ook voor visuele, auditieve, intra- en interpersoonlijke mogelijkheden. Indicatief voor deze evolutie zijn: de(h)erkenning van de verschillende mogelijkheden en talenten van leerlingen, geloof in de maakbaarheid van de mogelijkheden van leerlingen en het bieden van kansen voor leerlingen om verschillende mogelijkheden te ontwikkelen en te demonstreren.

2. Van toetsen naar multiple assessment

Traditioneel werd het leren van leerlingen enkel gemeten op basis van specifieke toetsen. Die bestonden uit gestructureerde en duidelijke vragen die binnen een korte tijdsperiode konden worden beantwoord en die enkel een beperkt deel van cognitieve kennis en vaardigheden dekten. Dergelijke 'paper- and pencil'-toetsen zijn echter maar één manier om gegevens te verzamelen over het leren van lerenden. Een bredere invulling van het evaluatieconcept is dan ook wenselijk. Daaronder verstaan we dat er meer en meer verscheiden informatie over leerlingen moet verzameld worden en dat de manier waarop en de doelen waarvoor deze informatie gebruikt wordt, moeten uitgebreid worden. Het repertoire aan leerstrategieën wordt bovendien geëvalueerd in communicatie met anderen en in toegepaste, realistische en cultureel diverse contexten.

3. Van geïsoleerde naar geïntegreerde assessment

Evaluatie moet enerzijds leerlingen feedback geven om het leerproces strategisch bij te sturen en anderzijds leerkrachten informeren over de noodzaak het onderwijsleerproces aan te passen om tegemoet te komen aan specifieke behoeftes van hun leerlingen (Rudman, 1989). Evaluatie helpt leerkrachten verwachtingen en standaarden te communiceren over leerlingen en prestaties. Kenmerkend voor recente assessmentbenaderingen is het feit dat men meer aandacht is gaan besteden aan de interactie tussen de evaluatie en het onderwijsleerproces in de klas en minder aan specifieke eigenschappen van de assessmenttechnieken op zich. Tegelijk met de beweging van beperkte toetsen naar multidimensioneel evalueren, is er dus sprake van een beweging naar integratie van evaluatie in het instructieproces in de klas. Het evaluatiegebeuren wordt niet louter meer aangezien als een afsluiting van een instructieperiode. Leren wordt immers gezien als een vloeiende beweging in en uit de verschillende taken en leerfasen. Ook evaluatie moet dus aangezien worden als een voortdurend dynamisch proces dat ook op zich bepaalde doelstellingen kan trachten te realiseren.

Tot zover de beschrijving van de evolutie van de traditionele toetscultuur naar een assessmentcultuur. We beschreven de belangrijkste evoluties en de kenmerken van beide culturen. In het tweede hoofdstuk van dit deel stellen we de vraag naar de stand van zaken in het Vlaamse onderwijs wat bovenstaande evolutie richting assessmentcultuur betreft. Alvorens daartoe over te gaan, staan we echter even stil bij de verhouding tussen beide evaluatieculturen.

2.3 Traditioneel en alternatief: tegengesteld of complementair?

De beschrijving van traditionele en hedendaagse evaluatiebenaderingen stelt enkele bemerkingen aan de orde. Daartoe richten we onze blik eerst op onze eigen concrete onderwijservaringen.

Even stilstaan... 1.5: Traditionele en hedendaagse onderwijsbenaderingen

Hoe verhouden de traditionele en hedendaagse onderwijs- en evaluatiebenaderingen zich in de concrete Vlaamse onderwijspraktijk? Een aantal vragen.

- ✓ Wat vindt u van de benaming 'traditionele' en 'hedendaagse' onderwijs- en evaluatiebenadering? Herkent u nog sporen van de 'traditionele' benaderingen in het hedendaagse onderwijs? En verder: herinnert u zich sporen van de 'hedendaagse' onderwijsbenadering in het onderwijs dat u destijds genoot?
- ✓ Herkent u in uw eigen onderwijspraktijk een evolutie van een traditionele naar een hedendaagse onderwijs- en evaluatiebenadering? Kan u die evolutie illustreren met een aantal voorbeelden.
- ✓ Gaat u steeds uit van dezelfde benadering of is er eerder sprake van een afwisseling tussen beide? Wanneer leunt uw onderwijspraktijk aan bij de alternatieve onderwijs- en evaluatiebenaderingen en wanneer bij de traditionele?

Boven werden de zogenaamde traditionele en hedendaagse evaluatiebenaderingen als het ware tegenover mekaar geplaatst. Wellicht is het beter beide als complementair te beschouwen. Afhankelijk van de doelen die men voor ogen heeft, zal immers één van beide benaderingen het meest gepast zijn. Er bestaat dus niet één bepaalde correcte (en meest geschikte) manier om leerlingen te onderwijzen en te evalueren. Ook al besteden we in dit onderzoek uitgebreid aandacht aan de zogenaamde alternatieve evaluatievormen, toch willen we benadrukken dat dit niet impliceert dat elke evaluatie van dit type moet zijn en dat het gebruik van traditionele toetsmethoden in het geheel moet verworpen worden (Herman et al., 1992; Arter & Bond, 1996; Stiggins, 1996). In volgende alinea's duiden we de complementariteit van de verschillende evaluatievormen.

2.3.1 Een verscheidenheid aan doelen vergt een verscheidenheid aan evaluatievormen

Doorheen dit handboek benadrukken we dat zogenaamde alternatieve evaluatie (zoals vaardigheidsevaluatie en authentieke evaluatie) interessante mogelijkheden biedt om complexe en probleemoplossende vaardigheden van leerlingen te evalueren. De traditionele evaluatievormen zouden daarentegen sterke tekorten hebben en inferieur zijn ten opzichte van de alternatieve evaluatievormen. Deze uitspraak klopt echter slechts gedeeltelijk. Het is correcter de situatie als volgt voor te stellen: 1. de ene evaluatievorm is in wezen niet superieur ten opzichte van de andere en 2. beide groepen evaluatievormen zijn zeer doeltreffend in het vervullen van andere specifieke doelstellingen. De alternatieve en de traditionele gestandaardiseerde evaluatievormen moeten met andere woorden complementair aangewend worden naargelang de doelstellingen die men wenst te evalueren.

De alternatieve evaluatievormen die in dit boek zullen worden voorgesteld, vormen een antwoord op de evaluatiebehoefte indien de onderwijsdoelen vereisen dat leerlingen complexe denkstrategieën demonstreren, specifieke taken moeten kunnen volbrengen, producten moeten kunnen creëren en dergelijke (Stiggins, 1996). Daartegenover staat bijvoorbeeld dat toetsen met meerkeuzevragen wellicht beter geschikt zijn om te evalueren in welke mate leerlingen essentiële feiten en concepten beheersen. Een gebalanceerd curriculum vereist dan ook een even uitgewogen geheel van evaluatietechnieken. De recente tendens die erin bestaat via onderwijsdoelen ook (metacognitieve) vaardigheden te benadrukken, zou logischerwijze dan ook moeten resulteren in een shift naar de zogenaamde alternatieve evaluatievormen. Vandaar de aandacht voor deze evaluatievormen in het verdere verloop van dit handboek.

2.3.2 Hoe nieuw is hedendaags?

Een bijkomende opmerking is hier aan de orde. Hoewel de alternatieve evaluatievormen de naam hebben nieuw te zijn, sluiten ze aan bij elementen waaraan sommige leerkrachten reeds jaren aandacht besteden. Goede leerkrachten richten hun blik steeds op het instructieproces -hoe verloopt een les, wie heeft er problemen, wie let er op, hoe doet een bepaalde groep het- om informatie te verzamelen over hoe de instructie en het curriculum aan te passen. Zij gebruiken tevens een rijke bron aan evaluatiegegevens om het leren van leerlingen te waarderen. Nieuw is bij deze leerkrachten dat de zogenaamde alternatieve evaluatievormen expliciet en formeel maken wat vroeger impliciet en informeel gebeurde.

Deze opmerking neemt evenwel niet weg dat er verscheidene indicaties zijn om de conclusie te rechtvaardigen dat een groot aantal leerkrachten nog een lange weg moet afleggen eer er sprake kan zijn van een assessmentcultuur. Deze indicaties worden in het volgende hoofdstuk toegelicht, samen met andere gegevens die relevant zijn om de huidige stand van zaken in Vlaanderen te schetsen.

Hoofdstuk 2: Assessmentcultuur - een stand van zaken

Zoals we hoger reeds stelden, worden de alternatieve onderwijs- en evaluatiebenaderingen uitvoerig in de literatuur beschreven. Deze vaststelling hoeft echter niet te betekenen dat de zogenaamde assessmentcultuur ook in de dagelijkse onderwijspraktijk reeds frequent is terug te vinden. De vraag is dus of de assessmentcultuur enkel op papier bestaat. Om de noodzaak van het ontwikkelen van alternatieve evaluatiemethoden na te gaan, willen we stilstaan bij de vraag of de huidige Vlaamse onderwijspraktijk eerder aansluit bij de traditionele dan wel bij de alternatieve onderwijsbenadering. In de volgende paragrafen gaan we daartoe eerst kort na of -en zo ja welke- regelgevingen, initiatieven en werkwijzen van de centrale overheid kenmerkend zijn voor de evolutie naar een assessmentcultuur. Daarna bestuderen we de concrete evaluatiepraktijk in Vlaamse secundaire scholen.

1 Evaluatiebenaderingen op macroniveau

Gelet op de belangrijke rol van het onderwijsbeleid inzake het in stand houden, tegengaan en veranderen van bepaalde opvattingen en praktijken in het onderwijsveld bekijken we eerst hoe ver het Vlaamse onderwijsbeleid zelf de kanteling naar een assessmentcultuur reeds heeft door gemaakt. Om de Vlaamse evaluatiebenaderingen op macroniveau te bestuderen, richten we ons op drie beleidsaspecten: eindtermen en ontwikkelingsdoelen, regelgeving en kwaliteitsbewaking, en basiscompetenties van leerkrachten.

1.1 Eindtermen en ontwikkelingsdoelen

In Vlaanderen werden in de eerste graad van het secundair onderwijs op 1 september 1997 het gebruik van eindtermen en ontwikkelingsdoelen ingevoerd² (Decreet 24-06-1996, Ministerie van de Vlaamse Gemeenschap, 1996a). Eerst geven we enige toelichting bij beide vormen van onderwijsdoelen, daarna staan we stil bij hun impact op het evaluatiegebeuren in Vlaamse scholen.

1.1.1 Definities en achtergronden

Eindtermen zijn minimumdoelen die de overheid noodzakelijk en bereikbaar acht voor een bepaalde leerlingengroep in Vlaanderen. Eindtermen zijn enerzijds een minimum aan kennis, inzicht en vaardigheden die alle leerlingen van de leerlingengroep verwerven tijdens het leer-

² De vakgebonden en vakoverschrijdende eindtermen van de tweede graad worden wellicht ingevoerd vanaf het schooljaar 2002-2003. De invoering van de eindtermen van de derde graad zal daar onmiddellijk op volgen.

proces en anderzijds een minimum aan attitudes die de school nastreeft bij de leerlingen. Eindtermen worden vastgelegd per graad en per onderwijsvorm. Er zijn vakgebonden en vakoverschrijdende eindtermen. Vakgebonden eindtermen dienen door de school ten aanzien van de leerlingen worden bereikt. Vakoverschrijdende eindtermen zijn minimumdoelen die niet specifiek behoren tot een vakgebied, maar onder meer door middel van meerdere vakken of onderwijsprojecten kunnen worden gerealiseerd; zij dienen enkel te worden nagestreefd. De vakoverschrijdende eindtermen omvatten vijf domeinen: lerenleren, sociale vaardigheden, opvoeding tot burgerzin, gezondheidseducatie en milieueducatie. Ook attitudinale eindtermen dienen enkel te worden nagestreefd.

Omwille van het specifieke karakter van het kleuteronderwijs en de B-stroom van het secundair onderwijs worden de doelstellingen daar **ontwikkelingsdoelen** genoemd. Ontwikkelingsdoelen zijn minimumdoelen op het vlak van kennis, inzicht, vaardigheden en attitudes die de onderwijsoverheid wenselijk acht voor een bepaalde leerlingenpopulatie. Voor ontwikkelingsdoelen geldt enkel een inspanningsverplichting, ze moeten uitdrukkelijk worden nagestreefd.

In de volgende figuur is een selectie van de eindtermen voor de eerste graad van het secundair onderwijs A-stroom samengebracht.

Figuur 1.8: Illustratie van vakgebonden en vakoverschrijdende eindtermen

Vakgebonden eindtermen wiskunde

De leerlingen...

- ✓ kunnen zich vanuit diverse vlakke weergaven een beeld vormen van een eenvoudige ruimtelijke figuur met behulp van allerlei concreet materiaal;
- ✓ passen probleemoplossende vaardigheden toe, zoals: het herformuleren van een opgave en het maken van een goede schets of een aangepast schema (vaardigheid);
- ✓ ontwikkelen bij het aanpakken van problemen zelfstandigheid en doorzettingsvermogen (attitude).

Vakoverschrijdende eindtermen 'leren leren'

De leerlingen ...

- ✓ kunnen informatiebronnen adequaat raadplegen;
- ✓ kunnen hun werktijd plannen en het nodige materiaal selecteren en ordenen;
- ✓ zijn bereid zelf naar oplossingen te zoeken en durven leer- en studieproblemen signaleren en uitleg of hulp vragen.

Eindtermen (en ontwikkelingsdoelen) zijn geen leerplannen. De organisatoren van het onderwijs zullen hun leerplannen baseren op deze eindtermen, met de mogelijkheid om dit minimum verder aan te vullen volgens hun eigen onderwijsproject. Vooral onder impuls van de koepelorganisaties worden de eindtermen bekeken als een minimum voor leerplanmakers (DVO, 1998). Vakinhoudelijke eindtermen moeten bereikt worden. Attitudes en vakoverschrijdende doelen moeten worden nagestreefd, wat betekent dat daar een inspanningsverplichting geldt. Scholen

zullen die doelstellingen dus in hun opvoedingsproject moeten aanbieden zonder ze te moeten realiseren.

1.1.2 Eindtermen en evaluatie

De eindtermen en ontwikkelingsdoelen geven het secundair onderwijs een impuls om op een ruimere manier naar het presteren van leerlingen te kijken. Uit de volgende alinea's kan blijken dat de eindtermen en ontwikkelingsdoelen aansluiten bij de hoger beschreven hedendaagse onderwijsbenaderingen en de assessmentcultuur (DVO, 1998; Klasse, 1997; Klasse, 1998).

“De eindtermen en ontwikkelingsdoelen steunen op een aantal basiselementen in de ontwikkeling van leerlingen. Zo zouden de leerlingen moeten beschikken over een positief zelfbeeld en moeten ze gemotiveerd zijn, zelf initiatief nemen, kunnen communiceren en samenwerken, zelfstandigheid aan de dag leggen en creatief en probleemoplossend omgaan met de hen omringende wereld. Dat is de werkelijkheid waarin het kind functioneert. De leerling leert de werkelijkheid begrijpen, wordt vaardig en ontwikkelt een positieve houding.

Telkens benadrukken de eindtermen het belang van sociaal functioneren en zelfstandig verwerken van informatie. Vooral bij de doelstellingen voor attitudes vinden we verwijzingen naar de interculturele en emancipatorische functie van het onderwijs terug. De eindtermen en ontwikkelingsdoelen hebben aandacht voor een brede vorming van de totale persoonlijkheid van het kind. Dat vereist niet alleen aandacht voor de cognitieve ontwikkeling, maar ook voor de psycho-motorische, dynamisch-affectieve en sociale ontwikkeling. Voorts leggen ze steeds de nadruk op het leren als actief en constructief proces. Bovendien moeten zoveel mogelijk kinderen zich een brede basisvorming eigen kunnen maken. Dit houdt in dat leerkrachten problemen tijdig moeten signaleren en remediëren. De eindtermen maken tevens een horizontale en verticale samenhang tussen vakken en leerjaren mogelijk. Leergebieden zijn dus geen aparte eilanden meer.”

Hoger stelden we dat een bepaalde visie op onderwijs (met bepaalde specifieke doelstellingen) zich moet vertalen in een aangepaste evaluatiebenadering. De na te streven doelstellingen (c.q. eindtermen) worden door de overheid verplicht als minimum te bereiken vooropgesteld. Dit impliceert dat de aard van deze doelstellingen tevens vereist dat aangepaste evaluatiemethoden worden aangewend. Dat brengt ons bij de vraag of de overheid ook wat de evaluatie van deze doelstellingen betreft bepaalde verplichtingen oplegt.

1.2 Evalueren: regelgeving en kwaliteitsbewaking

De onderwijsinspectie heeft onder meer als opdracht via schooldoorlichting te controleren of de school een voldoende studiepeil realiseert en voldoende inspanningen levert om de eindtermen

te bereiken. Maar hoe de school intern de leerlingen evalueert, behoort tot het domein van de pedagogische vrijheid. Dit betekent evenwel niet dat de inspectie geen aandacht heeft voor het evaluatiesysteem. Het is een belangrijke kwaliteitsindicator voor het onderwijsproces. De school moet haar evaluatiesysteem daartoe op elk ogenblik kunnen verantwoorden. In essentie komt het erop neer dat de leerkrachten evalueren op basis van een degelijk op school- of klasniveau uitgewerkt evaluatiesysteem. De school is bijvoorbeeld niet verplicht de in Vlaanderen sterk ingeburgerde trimesteriële examens te organiseren, maar ze moet wel een evaluatiesysteem hebben. Een uitzondering is de geïntegreerde proef die verplicht is in sommige leerjaren en onderwijsvormen van het secundair onderwijs³.

Scholen genieten voor de evaluatie van leerlingen, omwille van de pedagogische vrijheid, dus een grote autonomie. De overheid schrijft dus niet rechtstreeks voor op welke manier scholen tewerk moeten gaan bij het evalueren van de eindtermen en de ontwikkelingsdoelen.

Dat neemt echter niet weg dat de specifieke inhoud van de doelstellingen ook van invloed zal zijn op de evaluatiepraktijk. Bepaalde doelstellingen vereisen immers aangepaste evaluatiemethodes. Een greep uit de eindtermen kan dit illustreren:

- *Correct kunnen schrijven* is tegelijkertijd een kwestie van kennis van de spellingsregels, van vaardigheid in het schrijven als van houding om zonder fouten te willen schrijven.
- *Veiligheid in het verkeer* veronderstelt niet alleen de kennis van de regels maar ook de bereidheid om zich aan die regels te houden.
- *De leerlingen kunnen in een groepsdiscussie hun mening weergeven, handhaven en bijsturen.*

Om de kennis, inzichten, onderwijsattitudes en vaardigheden die noodzakelijk moeten bereikt worden (om van een realisatie van deze eindtermen te kunnen spreken) goed te evalueren, kan het niet volstaan uit te gaan van traditionele evaluatiebenaderingen en -methodes. In die zin zijn de eindtermen dus ook een stimulans en impuls (in de zin van indirecte verplichting) om leerkrachten ertoe te brengen ruimer en anders te evalueren.

1.3 Basiscompetenties van de leerkracht

In 1998 werd door de Vlaamse regering het decreet betreffende de beroepsprofielen van de leraren goedgekeurd (Decreet 29 september 1998, Ministerie van de Vlaamse Gemeenschap (1999)). Van deze beroepsprofielen worden basiscompetenties afgeleid waarop opleidingsinstellingen (hogescholen en universiteiten) hun opleidingsprogramma's voor toekomstige onderwijzers en leerkrachten zullen baseren. Op die manier zal het decreet mede richtinggevend zijn voor toekomstige ontwikkelingen aangaande onderwijs en evaluatie in Vlaamse scholen.

³ Gebaseerd op een toelichting van de toenmalige inspecteur-coördinator secundair onderwijs (Chris Blancquaert) in Klasse voor Leerkrachten (1998).

In volgende alinea's gaan we na wat de Vlaamse overheid van leerkrachten verwacht in verband met het evalueren van leerlingen. Ook op het vlak van de evaluatie stelt men immers enkele competenties voorop. Gelet op de nadruk die we reeds herhaaldelijk legden op de samenhang van onderwijsvisie, doelstellingen, curriculum en evaluatie richten we de aandacht eerst kort op de onderwijsprincipes die aan de basis liggen van de beroepsprofielen en basiscompetenties. De uitgangspunten leren het volgende over de achterliggende onderwijsvisie (Ministerie van de Vlaamse Gemeenschap, 1999, p. 6-7):

“Naast het waarborgen van een brede en harmonische vorming gaat het om een aantal onderwijsprincipes, die alle nadruk leggen op actief en constructief leren binnen een krachtige leeromgeving. Deze principes worden geformuleerd onder de vorm van tegenstellingen in die zin dat men in het onderwijs een verschuiving wil realiseren naar volgende nagestreefde polen: algemene vorming, toegepaste kennis, harmonische vorming, totaliteitsonderwijs, exemplarisch onderricht en beklijvend leren.”

De vraag is of dergelijke onderwijsvisie resulteert in bepaalde verwachtingen en competenties die al dan niet getuigen van een assessmentcultuur. Illustratief beschrijven we een selectie van wat in de beroepsprofielen en basiscompetenties voor leerkrachten vooropgesteld wordt aangaande het evalueren van leerlingen.

Er wordt van de leerkracht verwacht dat hij (Ministerie van de Vlaamse Gemeenschap, 1999, p. 93)⁴:

- de evaluatie en observatie kan voorbereiden: “De leraar kan in overleg doelstellingen, valide vragen, taken en opdrachten in diverse vormen kiezen en eventueel opstellen. Wat het observeren betreft, kan de leraar in overleg met collega's observatie-instrumenten kiezen en desgevallend opstellen. Hij kan ook de functie van een specifiek evaluatiemoment bepalen. Met ondersteuning kan de leraar beoordelingscriteria bepalen. De benodigde kennis omvat evaluatietechnieken en -instrumenten, (leerling)volgsystemen en een visie op evaluatie met het oog op bijsturing en remediëring.”
- vaardig is in het observeren en evalueren van het proces en het product: “De leraar kan op systematische wijze gegevens verzamelen via toetsen, observaties, zelfevaluatiegegevens van de lerende en gesprekken. Verder kan de leraar de vorderingen en prestaties correct interpreteren en beoordelen. De leraar kan, met ondersteuning van het team, adviezen geven over de voortgang van de lerenden in hun schoolloopbaan. De leraar kan activiteiten voor remediëring voorstellen. Hij kan ook studieresultaten rapporteren en bespreken. De leraar kan het eigen didactisch handelen toetsen aan evaluatiegegevens. De benodigde kennis omvat evaluatietechnieken en -

⁴ Deze vaardigheden worden gecategoriseerd onder de typefunctie 1: De leraar als begeleider van leer- en ontwikkelingsprocessen.

instrumenten, (leerling)volgsystemen, een visie op evaluatie in functie van bijsturing en remediëring, de techniek van foutenanalyse en van zelfevaluatie."

Uit deze beschrijving kan blijken dat ook in de beroepsprofielen en basiscompetenties de kante-ling van een traditionele toetscultuur naar een assessmentcultuur bepleit wordt. Getuige hiervan onder andere de nadruk op de verscheidenheid aan evaluatiemateriaal en de verschillende func-ties van evaluatie.

1.4 Afrondend

Bovenstaande beschrijving van een aantal onderwijsaspecten is indicatief voor het feit dat de overgang naar een assessmentcultuur ook op macroniveau bepleit wordt. Omwille van de peda-gogische vrijheid genieten scholen echter een grote autonomie omtrent de evaluatie van leerlin-gen. Een beschrijving van de regelgevingen en onderwijs- en evaluatievisies op macroniveau hoeft dus niet te impliceren dat de concrete evaluatiecultuur in de Vlaamse scholen hier effectief bij aansluit. Het verdient dan ook aanbeveling de blik eveneens te richten op de effectieve eva-luatiepraktijk in Vlaamse scholen.

2 Evaluatie in Vlaamse scholen

Onder volgende titels beschrijven we onderzoek dat de huidige Vlaamse evaluatiepraktijk be-studeert en trachten we een verklaring te geven voor de discrepantie tussen de in de literatuur beschreven en de reële evaluatiebenaderingen.

Zoals boven reeds toegelicht werd, kadert het ontwikkelen van alternatieve evaluatievormen in een ruimer onderzoeksproject. Een wezenlijk onderdeel hiervan betreft een schriftelijke bevraging van leerkrachten en directeurs over de evaluatiepraktijk in hun klas/school (Verhoeven et al., 2001). De meest relevante resultaten van dit onderzoek worden hier kort weergegeven. In-dien beschikbaar bekijken we eveneens of andere onderzoeksresultaten tot gelijkaardige conclu-sies leiden.

2.1 Kenmerken van de Vlaamse evaluatiepraktijk

Om de evaluatie in Vlaamse scholen te beschrijven, rapporteren we kort de onderzoeksresulta-ten aangaande drie evaluatieaspecten: de gebruikte evaluatie-instrumenten, de functies van evaluatie en de inhoud van de evaluaties.

2.1.1 Gebruik van evaluatie-instrumenten

In een vragenlijst werden leerkrachten secundair onderwijs bevraagd over welke evaluatie-instrumenten ze gebruiken. Uit de topdrie antwoorden blijkt dat bijna alle leerkrachten schrifte-lijk evalueren (toetsen, schriftelijk examen en huistaken). Er wordt dan ook geconcludeerd dat leerkrachten nog vaak -op formele wijze- traditionele evaluatie-instrumenten hanteren. Deze constatering maakt dat er zeker nog sprake is van kenmerken van de zogenaamde toetscultuur.

Leerkrachten geven wel aan dat er veel gebruik wordt gemaakt van (ongestructureerde) obser-vaties om de leerlingenprestaties en -attitudes te evalueren. Bovendien geeft meer dan de helft van de leerkrachten aan dat leerlingen zichzelf kunnen evalueren en dat ze gebruik maken van discussies. Leerkrachten erkennen wel dat leerlingen in staat zijn zichzelf te evalueren maar gaan zeer zelden over tot het gebruik van zelfevaluatie-instrumenten.

Naargelang van het soort vak dat de leerkracht geeft, worden er statistisch significante ver-schillen gerapporteerd ten aanzien van het gebruik van bepaalde evaluatie-instrumenten. Leer-krachten algemene en technische vakken maken significant meer gebruik van schriftelijke toet-sen dan praktijkleerkrachten. En verder: leerkrachten algemene vakken maken meer gebruik van schriftelijke examens en huistaken dan hun collega's die technische en praktijkvakken geven. Leerkrachten praktische vakken maken dan weer significant meer gebruik van permanente eva-luatie en zelfevaluatie dan de leerkrachten algemene en technische vakken. Op zich zijn dit te verwachten resultaten. Ook tussen de graden worden significante verschillen gevonden. Een merkwaardige vaststelling is dat hoe hoger de betreffende graad, hoe minder leerkrachten aan-geven dat ze de leerlingen zichzelf laten evalueren.

Er werd leerkrachten eveneens gevraagd in welke mate ze alternatieve evaluatie-instrumenten gebruiken. Op basis van de onderzoeksgegevens moet geconcludeerd worden dat alternatieve evaluatiemethoden (zoals portfolio's, peerevaluatie, leerlingprofiel-beschrijvingen, logboeken) slechts zeer beperkt gebruikt worden in het secundair onderwijs. Leerkrachten gebruiken voor-namelijk traditionele evaluatie-instrumenten. Verscheidene leerkrachten percipiëren wel dat ze alternatief evalueren maar dit betreft dan zaken als een openboek-toets en een computerevalua-tie en dus niet noodzakelijk wat wij zullen verstaan onder alternatieve evaluatievormen. Wat het gebruik van evaluatie-instrumenten betreft is er dus nog geen sprake van een assessmentcultuur.

2.1.2 De functies van evalueren

Een belangrijk verschilpunt tussen de toets- en de assessmentcultuur betreft de functies en doe-len die vooropgesteld worden bij het evalueren van leerlingen. Op basis van de functies die in de literatuur beschreven worden, bevroegen Verhoeven et al. (2001) drie functies, namelijk: beoordelen, bijsturen van het leerproces en begeleiden van de leerling.

1. Beoordelen

De scores op de verschillende aspecten binnen de vragenlijst leren dat leerkrachten in het algemeen belang hechten aan elk van de drie functies van evalueren. Toch moet men constateren dat leerkrachten nog steeds 'het beoordelen van leerlingen' als de meest belangrijke functie van evalueren aanzien. Ook Struyf (2000) en Janssens et al. (2000) stellen dat de functie 'oordelen of de leerlingen de leerstof beheersen' als de meest belangrijke naar voren wordt geschoven. Zij concluderen dat de evaluatiepraktijk van elke leerkracht min of meer als summatief moet gekarakteriseerd worden. Bijna alle leerkrachten geven aan dat ze de toetsen voorzien van een cijfermatige beoordeling die (gedeeltelijk) meetelt in de beoordeling op het einde van het schooljaar. Leerkrachten geven nooit een schriftelijke toets zonder een beoordeling in punten.

Uit de resultaten van de bevraging wordt tevens geconcludeerd dat leerkrachten in de eerste graad en leerkrachten uit ASO-richtingen de beoordelende functie respectievelijk belangrijker vinden dan hun collega's uit hogere graden en dan hun collega's uit het TSO en BSO.

2. Begeleiden van de leerling

De functie 'begeleiden van leerlingen' blijkt als iets minder belangrijk beschouwd te worden dan beoordelen. In het onderzoek van Struyf (2000) blijken verwante functies eveneens hoog te scoren wat hun relatieve belang betreft. In concreto geldt dit voor volgende functies: toetsen dienen om leerlingen informatie te geven zodat ze weten wat ze fout doen en wat ze al dan niet kunnen en kennen, en toetsen dienen om na te gaan welke fouten leerlingen nog maken zodat hier in de lessen kan op ingegaan worden. Zowel vanuit het standpunt van de leerling als vanuit het standpunt van de leerkracht wordt deze evaluatiefunctie dus belangrijk geacht.

3. Bijsturen van het leerproces

Deze functie wordt door de leerkrachten over de ganse lijn als de minst belangrijke beschouwd. In vergelijking met andere doelen wordt in het onderzoek van Struyf (2000) de functie 'toetsen zijn een middel om na te gaan of de les wel goed gegeven is' als de minst belangrijke omschreven.

Concluderend: Vlaamse leerkrachten vinden de summatieve functie (beoordelen) van evaluatie belangrijker dan de (meer) formatieve functies (zoals begeleiding van leerlingen en bijsturen van het leerproces). We kunnen dus stellen –uitgaande van de functies die leerkrachten aan evaluatie toeschrijven– dat de toetscultuur in Vlaamse scholen de assessmentcultuur overheerst.

2.1.3 Welke doelstellingen worden geëvalueerd

De vragenlijst bij het onderzoek van Verhoeven et al. (2001) bevroeg eveneens in welke mate de leerkrachten volgende soorten doelstellingen evalueren: kennis, inzicht, vaardigheden en attitudes. De leerkrachten geven aan dat ze het meest nagaan of leerlingen nauwkeurig kunnen

werken en of ze verbanden kunnen leggen binnen de leerstof. Het minst vaak zouden ze letterlijk bepaalde leerstofonderdelen evalueren. Kennis van begrippen en feiten wordt evenwel nog vaak geëvalueerd. De onderzoekers merken hierbij op dat leerkrachten ten aanzien van het letterlijk evalueren sociaal wenselijk kunnen hebben geantwoord.

2.1.4 De samenhang tussen frequentie, gebruik en functie

Uit de analyses blijkt dat leerkrachten summatieve evaluatie belangrijker vinden dan formatieve evaluaties. Dit hangt wellicht samen met het belang dat aan de beoordelende functie van evaluatie wordt gehecht. De analyses in dit onderzoek tonen een beperkte (doch statistisch significante) correlatie tussen het belang van 'beoordelen' als functie van evaluatie en het gebruik van schriftelijke toetsen en examens. De functie 'bijsturen van het leerproces' kent een grotere samenhang met het frequent gebruik van volgende evaluatie-instrumenten: demonstreren, discussie en zelfevaluatie. De functie 'begeleiden van leerlingen' gaat op haar beurt gepaard met een frequenter gebruik van informele observaties.

Uit deze samenhang blijkt het belang van de visie van leerkrachten op evaluatie en de functies ervan. De functies die men aan evalueren toeschrijft, zijn immers bepalend voor welke evaluatie-instrumenten leerkrachten gebruiken.

2.2 Concluderend: bestaat de assessmentcultuur vooralsnog enkel op papier?

Op basis van de gegevens over de eigenlijke evaluatiepraktijk in Vlaamse scholen kunnen we concluderen dat de zogenaamde toetscultuur nog steeds dominant is. Evaluatie gebeurt meestal aan de hand van formele schriftelijke toetsen en examens. We schreven dat het gebruik van deze evaluatie-instrumenten samenhangt met het belang dat gehecht wordt aan de beoordelende functie van evaluatie. Samen met het feit dat leerkrachten nog vaak de kennis van begrippen en feiten evalueren, rechtvaardigt dit de conclusie dat de Vlaamse evaluatiepraktijk voornamelijk gekenmerkt wordt door de toetscultuur. In het kwantitatieve luik van dit onderzoeksproject kwam men tot volgende conclusie (Verhoeven et al., 2001):

“De evolutie van toetscultuur naar een evaluatiecultuur die in de literatuur wordt beschreven, blijkt slechts in beperkte mate aanwezig in de Vlaamse klaspraktijk. 18% van de leerkrachten evalueert zeer frequent procesgericht. Nog steeds wordt het evalueren gezien als een beoordelend moment binnen de schoolloopbaan van de leerlingen. De andere functies van evalueren, zoals de begeleidings- of bijsturingsfunctie, zijn van secundair belang. Het moment van evalueren is een individueel moment. Samenwerken tijdens de evaluatie wordt weinig tot niet getolereerd omdat men de leerling wil beoordelen. Het gebruik van alternatieve evaluatie-instrumenten is nog miniem. Leerkrachten baseren zich vaak nog op de meer traditionele evaluatie-instrumenten. Sommige leer-

krachten geven geen nabespreking na een evaluatiemoment. Het geven van een cijfer met een korte commentaar is een feit binnen het secundair onderwijs."

Het vermoeden bestaat dat de toetscultuur nog meer van kracht is in het Vlaamse onderwijs dan uit het bevragen van leerkrachten blijkt. Het bevragen van opinies van leerkrachten aan de hand van vragenlijsten is immers wellicht niet de meest geschikte methode om een accuraat beeld te krijgen van de Vlaamse evaluatiepraktijk. Leerkrachten kunnen wel aangeven bepaalde functies belangrijk te vinden en bepaalde inhoud te evalueren; de vraag blijft echter in welke mate dit vertaald wordt in de praktijk. Hierbij lijken ons twee interpretaties mogelijk. Ten eerste: vele leerkrachten zijn van mening dat ze aanleunen bij de zogenaamde assessmentcultuur terwijl ze daar objectief nog een eind van verwijderd zijn. Ten tweede: de leerkrachten zijn zich wel degelijk bewust van wat in de onderwijsliteratuur bepleit wordt en ze geven bij de bevraging aan volgens die principes tewerk te gaan (in deze interpretatie is er dus sprake van sociale wenselijkheid). Beide interpretaties impliceren echter dat de assessmentcultuur vooralsnog inderdaad voornamelijk op papier bestaat.

3 Geen eenduidig beeld: verschillende evaluatiepraktijken

Bovenstaande onderzoeksgegevens trachten een algemeen beeld te geven van de evaluatiepraktijk in Vlaamse scholen. Het is echter niet zo dat leerkrachten die in een bepaalde graad een bepaald vak onderwijzen allemaal stereotypisch tegemoet komen aan het beschreven beeld. Dit zou bijvoorbeeld betekenen dat alle leerkrachten geschiedenis er dezelfde evaluatiepraktijk op na zouden houden. Wellicht is het beter binnen het Vlaamse onderwijs een aantal bestaande evaluatiepraktijken te onderscheiden. In de volgend paragraaf rapporteren we kort de drie verschillende praktijken die door Struyf (2000) in lijn met deze gedachtegang onderscheiden worden. Op die manier willen we het huidige Vlaamse evaluatiebeleid via een andere invalshoek kenmerken. Aansluitend op de beschrijving van de verschillende evaluatiepraktijken richten we ons opnieuw op de vraag of de assessmentcultuur enkel op papier bestaat.

3.1 Beschrijving van te onderscheiden evaluatiepraktijken

Naargelang van de fase die of het deelaspect van evaluatie dat men onder de loep neemt, kunnen verschillende soorten evaluatiepraktijken beschreven worden. In het kader van het onderscheid tussen de toets- en de assessmentcultuur richten we ons specifiek op de laatste fase in het evaluatieproces, namelijk de terugkoppelingsfase. Op basis van wat leerkrachten na het verbeteren en beoordelen van evaluatiegegevens doen, kunnen we een indicatie krijgen van de doelen die leerkrachten vooropstellen bij het evalueren. Struyf (2000) maakt een onderscheid tussen drie evaluatiepraktijken, namelijk de beoordelingsgerichte, de oplossingsgerichte en de procesgerichte. Deze worden hier achtereenvolgens beschreven.

Beoordelingsgerichte leerkrachten hebben niet de gewoonte na de evaluatie (toets, examen,...) een follow-up activiteit te organiseren. Ze geven schriftelijke toetsen meestal na correctie en beoordeling aan de leerlingen terug zodat deze kunnen zien hoe ze op de toets gepresteerd hebben. Er volgt in de regel geen individuele noch collectieve bespreking, tenzij de meerderheid van de leerlingen iets niet begrepen heeft. De beoordelingsgerichte leerkrachten geven aan dat ze toetsbesprekingen als tijdverlies beschouwen. De leerlingen moeten zelf nagaan of ze hun fouten begrijpen en eventueel zelf het initiatief nemen om verduidelijking te vragen. De leerkracht zal enkel zelf iets ondernemen als de antwoorden op een toetsvraag in het geheel niet voldoen aan de vooropgestelde verwachtingen.

Oplossingsgerichte leerkrachten beweren na toetsafname of na verbetering telkens de juiste antwoorden of oplossingen op de toetsvragen te geven. Vaak is deze bespreking zeer kort. De leerkrachten erkennen dat overhoringen geen zin hebben als de foutieve antwoorden niet verbeterd worden. Een toetsbespreking is dan ook geen tijdverlies. Toetsen dienen niet alleen om te beoordelen, leerlingen moeten er ook iets uit leren. Aan de hand van de juiste antwoorden kunnen de leerlingen bij zichzelf nagaan of ze het juiste antwoord begrijpen en waar ze eventueel fouten gemaakt hebben. Enkel indien veel leerlingen de vraag zeer slecht oplossen, wordt ook verder toegelicht op welke manier men tot het juiste antwoord had kunnen komen.

Leerkrachten wier praktijk gekenmerkt wordt door de procesgerichte evaluatiepraktijk besteden tijdens de nabesprekingen meer aandacht aan het oplossingsproces. Het is de bedoeling dat leerlingen inzicht krijgen in hun oplossingsstrategie. Door het oplossingsproces te bespreken, wil men leerlingen de nodige leerruimte bieden om te reflecteren en om vragen te stellen.

Net zoals oplossingsgerichte leerkrachten beschouwen procesgerichte leerkrachten het als hun verantwoordelijkheid om na elke evaluatie (c.q. toets) een nabespreking te voorzien. Beide groepen leerkrachten vinden bovendien dat leerlingen hun deel in de verantwoordelijkheid voor hun leren moeten opnemen.

3.2 Concluderend: bestaat de assessmentcultuur enkel op papier?

Beoordelingsgerichte leerkrachten organiseren in de eerste plaats toetsen om een oordeel uit te spreken over het kennen en kunnen van de leerlingen. Omdat zij slechts in uitzonderlijke gevallen remediërende activiteiten opzetten, wordt hun evaluatiepraktijk gekarakteriseerd als niet-formatief en dus als kenmerkend voor de zogenaamde toetscultuur. Door een toetsbespreking te organiseren, kunnen nieuwe leerkansen ontstaan voor leerlingen. Deze kansen zijn het grootst wanneer de leerling actief met de leerstof bezig is en wanneer de correcte oplossingswijze blootgelegd wordt. Met andere woorden, deze kansen zijn het grootst bij procesgerichte leerkrachten. Hun evaluatiepraktijk wordt gekarakteriseerd als formatief en als kenmerkend voor de zogenaamde assessmentcultuur.

Dit brengt ons bij de vraag welke evaluatiepraktijk kenmerkend is voor de Vlaamse leerkracht. Uit de studie van Struyf (1998) blijkt dat de evaluatiepraktijk van 38% van de bevraagde leraren procesgericht is en dat de evaluatiepraktijk van 37% van de bevraagde leraren oplossingsgericht is. 25% van de bevraagde leerkrachten is beoordelingsgericht. Wat het terugkoppelingsaspect van evaluatie betreft, lijkt dit hoopgevend. Hiertegenover staat het feit dat alle leerkrachten in het onderzoek –ongeacht hun evaluatiepraktijk- aangeven op schriftelijke toetsen punten te geven (Struyf, 1998, p. 178). Er wordt dus telkens in de vorm van een cijfer een waardeoordeel uitgesproken zodat er in strikte zin eigenlijk aan de hand van schriftelijk toetsen geen sprake is van assessmentcultuur.

4 Een discrepantie tussen literatuur en praktijk?

Evaluatie heeft altijd een belangrijke rol gespeeld in het onderwijsleerproces. Dit belang weerspiegelt zich in de aandacht die evaluatiebenaderingen kregen en krijgen in onderwijsliteratuur. Evaluatie is bijvoorbeeld een verplicht hoofdstuk in didactische handboeken (zoals Lowyck & Verloop, 1995; Standaert & Troch, 1998; Janssens et al., 2000). Struyf (2000) schrijft dat de evaluatieliteratuur overwegend prescriptief van aard is. De nadruk ligt op de verschillende aspecten van evalueren en er wordt geschetst hoe leraren het best te werk kunnen gaan. Er worden richtlijnen en vuistregels gegeven over hoe goede evaluatie-instrumenten moeten worden opgesteld en men besteedt aandacht aan hetgeen waarop een leerkracht bij het beoordelen moet letten. Cruciale vraag blijft echter in welke mate dergelijke elementen een weerslag vinden in de onderwijspraktijk. Uit bovenstaande onderzoeksgegevens bleek die weerslag vooralsnog beperkt, zodat we kunnen spreken van een discrepantie tussen evaluatieliteratuur en de praktijk.

4.1 Indicaties voor de discrepantie

Uit haar literatuuronderzoek leert Struyf (2000, p.31) dat 'inzichten die de laatste twee decennia min of meer gemeengoed zijn geworden in de evaluatieliteratuur nog maar weinig invloed hebben gehad op de evaluatiepraktijk'. Afgezet tegen de boven beschreven (theoretische) assessmentcultuur zijn volgende kenmerken van de evaluatiepraktijk van leerkrachten indicatief voor de bestaande discrepantie tussen literatuur en praktijk:

- formatieve evaluatie neemt slechts een beperkte plaats in;
- evaluatiegegevens worden zelden voor diagnostische doeleinden gebruikt;
- evaluatie wordt niet gebruikt om tot interne differentiatie te komen;
- er is geen systematische verzameling van gegevens over de beginsituatie van leerlingen;
- de meeste nadruk ligt op 'cijfers geven';
- de kwaliteit van de door de leerkrachten geconstrueerde toetsen is vaak laag;
- alternatieve evaluatiemethoden vinden maar moeilijk ingang in de klaspraktijk;
- evaluatie wordt onvoldoende geïntegreerd in het onderwijsleerproces.

4.2 De discrepantie verklaard

Bovenstaande indicaties en onderzoeksresultaten tonen aan dat er voldoende elementen zijn om te concluderen dat er inderdaad een discrepantie is tussen de evaluatietheorie en de concrete klaspraktijken. Struyf (2000, p. 32) geeft een aantal hypothetische verklaringen voor deze vaststelling.

Een mogelijke verklaring is dat er in de evaluatieliteratuur onvoldoende rekening wordt gehouden met de voorwaarden waaraan moet worden voldaan, willen de beschreven inzichten toepasbaar zijn in de dagelijkse onderwijspraktijk. Summatieve evaluatie vergt bijvoorbeeld een beperktere tijds- en arbeidsintensiviteit dan de in de literatuur vooropgestelde evaluatiemethoden. De bijkomende belasting van formatieve evaluatie en de bijbehorende praktische problemen zouden leerkrachten ervan kunnen weerhouden om dergelijke evaluatievormen te implementeren. Het komt er dan ook op aan aangepaste werkcondities te realiseren.

Een alternatieve verklaring stelt dat leerkrachten slechts in beperkte mate op de hoogte zijn van de nieuwe theoretische evaluatie-inzichten en dat zij bovendien niet over de geschikte evaluatievaardigheden beschikken. Te denken valt aan het belang dat gehecht wordt aan metacognitieve vaardigheden en de consequenties hiervan voor de evaluatiemethode. Het gebrek aan kennis zou tevens kunnen verklaren waarom leerkrachten hun evaluatiepraktijk niet als problematisch ervaren. Indicatief hiervoor is tevens de mate waarin leerkrachten sociaal wenselijk antwoorden op bepaalde gewichtige evaluatiethema's. Leerkrachten geven immers zelf aan dat ze de verschillende functies van evaluatie belangrijk vinden, dat ze reproductie uit den boze vinden en dat ze aandacht besteden aan de evaluatie van attitudes. We zouden kunnen stellen dat Vlaamse leerkrachten vaak onvoldoende op de hoogte zijn van de theoretische achtergronden van dergelijke onderwerpen en om die reden geen problemen ervaren. Als er dan toch problemen gerapporteerd worden, ontbreekt de know-how om de praktijk te veranderen. Struyf (2000) legt in dit opzicht een belangrijke rol weg voor de lerarenopleidingen om leerkrachten beter te gaan voorbereiden op het concreet uitvoeren van evaluaties binnen een assessmentcultuur.

Terloops willen we nog opmerken dat bovenstaande kenmerken van de evaluatiepraktijk vaak nog aansluiten bij de doelen van de traditionele onderwijs- en evaluatiebenadering. Veel aandacht blijft gaan naar de beoordelende (c.q. selecterende) functie van evaluatie en minder naar de diagnosticerende en begeleidende functie. Dit uit zich uiteraard in andere evaluatieopvattingen en -handelingen. Leerkrachten die de selectie- en beoordelingsfunctie van evaluatie benadrukken, zullen minder de noodzaak onderschrijven om instructie en evaluatie te integreren (Struyf, 2000).

Het komt er dan ook op aan de onderwijspraktijk van leerkrachten te bestuderen, in concreto bestaande opvattingen en concrete handelingen. Enkel indien leerkrachten het nut en de meerwaarde van alternatieve evaluatiebenaderingen en evaluatievormen onderschrijven, heeft de implementatie van nieuwe evaluatiemethodes een kans op slagen. Struyf (2000) schrijft dat het

ter beschikking stellen van concrete hulpmiddelen en nascholing leerkrachten moet leren omgaan met de nieuwe, alternatieve evaluatiemethoden en met hun rol als 'diagnosticus'.

Deel

2

Op zoek naar goede evaluatie

INHOUD - DEEL 2

Op zoek naar goede evaluatie

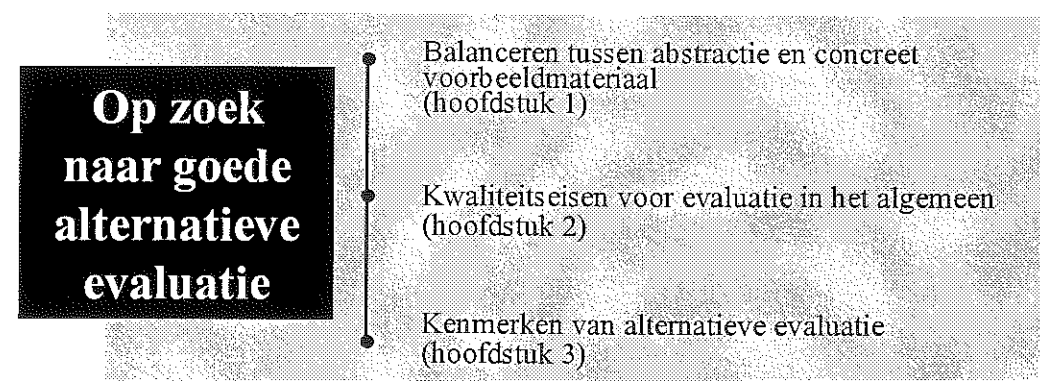
Inleidend	60
Hoofdstuk 1: Balanceren tussen abstractie en concreet voorbeeldmateriaal	61
1 Moet het voorbeeldmateriaal centraal staan?.....	61
1.1 Het belang van concretisering	62
1.2 Hoe tot concretisering komen?.....	62
2 Selectie van casussen.....	63
3 Selectie van doelstellingen.....	64
3.1 Waarom een taxonomie?.....	64
3.2 Inhoud van de taxonomie	65
4 Uitwerking van de casussen.....	68
4.1 Eerste graad – vakken Nederlands en Wiskunde	68
4.2 Tweede graad TSO - Elektriciteit/Elektronica	69
4.3 Derde graad BSO - Verzorging	69
Hoofdstuk 2: Kenmerken van goede evaluatie	71
1 Inleidend.....	71
2 Eisen van goede evaluatie	71
2.1 Doelmatigheid	72
2.2 Billijkheid.....	81
3 De evaluatie geëvalueerd	84
3.1 Toepassing van doelmatigheids- en billijkheidscriteria	85
3.2 Hoe de kwaliteit van een evaluatie bewaken?	86

Hoofdstuk 3: Kenmerken van alternatieve evaluatievormen	91
1 Inleidend.....	91
2 Kenmerken van alternatieve evaluatievormen.....	91
2.1 De evaluatie is aangepast aan de kenmerken van de leerlingen.....	93
2.2 De integratie van evaluatie met onderwijzen en leren.....	95
2.3 Betrokkenheid van leerkrachten en studenten.....	101
2.4 Kennisconstructie is het uitgangspunt, niet kennisreproductie.....	103
2.5 Gebruik van authentieke of levensechte prestatiesituaties.....	104
2.6 Concluderend: kenmerken van alternatieve evaluatievormen.....	104
3 Voldoen alternatieve evaluatievormen aan de eisen van goede evaluatie?.....	106
3.1 De doelmatigheid en billijkheid van alternatieve evaluatievormen.....	107
Samenvatting	109

Inleidend

Wat kenmerkt evaluatiemethoden die passen binnen een assessmentcultuur? Aan welke kwaliteitseisen moet een goede evaluatie voldoen? In welke zin kan je als leerkracht je eigen evaluaties bijsleutelen om tot betere resultaten te komen? Wanneer kan je spreken van een alternatieve evaluatievorm? Dergelijke vragen staan centraal in het tweede deel van dit handboek. We gaan met andere woorden op zoek naar de kenmerken van en de eisen aan goede en alternatieve evaluatie. Dit uit zich in volgende structuur.

Figuur 2.1: Structuur van deel 2



Het antwoord dat we zullen presenteren, leidt ons tot een aantal kwaliteitseisen die kunnen worden vooropgesteld ten aanzien van goede evaluatie in het algemeen en van zogenaamd alternatieve evaluatievormen in het bijzonder. We maken dus expliciet het onderscheid tussen beide groepen eisen aan evaluatie, tussen goede evaluatie en alternatieve evaluatie. Dit uit zich in een hoofdstuk 'kenmerken van goede evaluatie' en een hoofdstuk 'kenmerken van alternatieve evaluatievormen'. We willen hiermee benadrukken dat een evaluatie niet per definitie alternatief hoeft te zijn om als 'goed' te worden bestempeld. Alternatieve evaluatie gaat als het ware een stap verder dan goede evaluatie.

Alvorens deze hoofdbroek in de hoofdstukken 2 en 3 aan te snijden, staan we in het eerste hoofdstuk stil bij een aantal pragmatische en structurele bekommernissen bij het verdere verloop van dit handboek.

Hoofdstuk 1: Balanceren tussen abstractie en concreet voorbeeldmateriaal

In de eerste plaats stellen we ons in dit hoofdstuk de vraag in welke mate concreet voorbeeldmateriaal al dan niet centraal moet staan in een handboek dat (toekomstige) leerkrachten wil stimuleren en ondersteunen in hun zoektocht naar zo geschikt mogelijke evaluatiemethoden. Hierbij aansluitend worden een aantal casussen uitgewerkt die we doorheen de volgende delen en hoofdstukken zullen aanwenden om de verschillende evaluatieaspecten te illustreren.

1 Moet het voorbeeldmateriaal centraal staan?

Pas in de volgende hoofdstukken en delen staan we uitdrukkelijk stil bij verschillende aspecten van (alternatieve) evaluatie. Dit betekent dat we onze aandacht nog even niet richten op de ontwikkeling van concreet voorbeeldmateriaal van alternatieve evaluatievormen. Deze werkwijze willen we als volgt beargumenteren.

- ✓ Ondanks het feit dat het ontwikkelen van concrete voorbeelden een belangrijk doel is van dit handboek, moet het belang ervan in de meer omvattende opzet gerelativeerd worden. Als het omvattende doel is leerkrachten bewust te maken van bepaalde tekortkomingen van bestaande evaluatiepraktijken en hen een aanzet te bieden tot alternatieve praktijken, is het wellicht aangewezen hen eerder te voorzien van algemene aandachtspunten bij en criteria van een goed evaluatieproces dan van een batterij concrete voorbeeldinstrumenten. In dit opzicht zouden we de vergelijking kunnen maken met een kookboek: wat heeft men aan een foto van een heerlijk ogende schotel wanneer de ingrediënten en het bereidingsproces zelf niet toegelicht worden? Na het doornemen van dit handboek moeten leerkrachten zelf aan de slag kunnen. In deze context is het ontwikkelen van voorbeeldmateriaal dan ook eerder een hulpmiddel dan een doel op zich.
- ✓ Met de aandacht die we besteden aan het evaluatieproces en relevante criteria willen we leerkrachten tevens een spiegel voor ogen houden. Ons inziens wordt de noodzaak aan alternatieve evaluatievormen door leerkrachten beter gevoeld en onderschreven wanneer men vertrekt van essentiële kenmerken en kwaliteitsgaranties van evaluatie in het algemeen en van alternatieve evaluatie in het bijzonder. De huidige evaluatiepraktijk zal kritisch kunnen bekeken worden. Het louter presenteren van concrete evaluatieinstrumenten slaat daarentegen enkele essentiële stappen over. Op die manier wordt de toegankelijkheid van de boodschap te klein en zal de te overbruggen afstand voor vele leerkrachten te groot blijken.

Voorgaande elementen mogen evenwel niet impliceren dat we blijven steken in theoretische en abstracte beschrijvingen van kenmerken van goede evaluatie in het algemeen, van goede alternatieve evaluatievormen en van het evaluatieproces in het bijzonder. We moeten dus als het ware op zoek gaan naar een geschikte manier om de achterliggende principes toch te concretiseren.

1.1 Het belang van concretisering

In vorig punt stelden we dat het louter presenteren van concreet evaluatiemateriaal wellicht niet de meest geschikte werkwijze is om leerkrachten in te wijden in het gebruik van alternatieve evaluatievormen. Dit neemt niet weg dat we er -waar nodig- voor opteren de theoretische achtergrond in de volgende hoofdstukken te concretiseren. Volgende bedenkingen liggen aan de basis van deze keuze.

- ✓ Moeilijke abstracte begrippen worden maar inzichtelijk wanneer ze gehanteerd worden in een concrete context en geïllustreerd aan de hand van voorbeelden. Dit geldt des te meer wanneer men een doelgroep wenst te bereiken die niet vanzelfsprekend vertrouwd is met het gehanteerde begrippenkader.
- ✓ Concretisering maakt de voorgestelde elementen herkenbaarder voor leerkrachten en kan in die zin een facilitator zijn voor de transfer van aangehaalde aspecten naar de eigen klaspraktijk. Leerkrachten worden bovendien gestimuleerd om over de eigen evaluatiepraktijk te reflecteren.
- ✓ Doorheen de beschrijving van het evaluatieproces en van de criteria die goede alternatieve evaluatievormen kenmerken, wordt herhaaldelijk benadrukt dat de keuze van een evaluatie-instrument sterk afhankelijk is van contextuele kenmerken (zoals kenmerken van de leerlingen, te evalueren doelstellingen, doelen van de eigenlijke evaluatie, etc).

1.2 Hoe tot concretisering komen?

Bij het ontwikkelen van concreet voorbeeldmateriaal is het aangewezen de algemene lijn te volgen die in het theoretische kader beschreven en benadrukt werd. Dit betekent dat men bij het ontwikkelen en selecteren van evaluatiemethoden, evaluatie-instrumenten en evaluatietaken steeds moet uitgaan van de doelstellingen die men bij de leerlingen wil realiseren. Hiertoe stellen wij volgende stapsgewijze werkwijze voor:

1. Selecteren van een aantal prototypische groepen leerlingen waarvoor we mogelijke evaluatiemethoden kunnen ontwikkelen.

2. Voor elk van deze groepen leerlingen een selectie maken van bestaande doelstellingen (vakspecifiek en vakoverschrijdend) op basis van de eindtermen en bestaande leerplannen.
3. Uitwerken van voorbeeldmateriaal (mogelijke evaluatiemethoden en -instrumenten) dat de specificiteit van de prototypische doelgroepen en de aard van de verschillende doelstellingen in acht neemt.

De selectie van de casussen en het vooropstellen van doelstellingen worden in de volgende paragrafen verder toegelicht. Het uitwerken van concreet voorbeeldmateriaal zal vervolgens tijdens het verdere verloop van het handboek systematisch aan bod komen. Doorheen de beschrijving van de belangrijkste aspecten zullen wij met andere woorden regelmatig gebruik maken van casussen ter illustratie van de verscheidene theoretische en abstracte evaluatieaspecten.

2 Selectie van casussen

Het is uiteraard onmogelijk om voor heel het onderwijsveld specifiek en concreet voorbeeldmateriaal te ontwikkelen. Het komt er dan ook op aan een aantal prototypische voorbeelden te ontwikkelen. Een mogelijke werkwijze hiertoe bestaat erin een aantal -naar de onderwijsdoelen toe- specifieke doelgroepen te selecteren en het evaluatieproces voor deze doelgroepen te concretiseren. Wanneer deze specifieke doelgroepen goed gekozen worden, moet de uitwerking van het concrete evaluatieproces in het algemeen en van het evaluatiemateriaal in het bijzonder voor alle leerkrachten herkenbaar, leerrijk en inspirerend zijn.

Een voor de hand liggend onderscheid in het Vlaamse onderwijs betreft de opdeling in onderwijsvormen. In concreto denken we hierbij aan het algemeen secundair onderwijs (ASO), het technisch secundair onderwijs (TSO) en het beroepssecundair onderwijs (BSO). Volgende korte beschrijving van deze onderwijsvormen kan de fundamentele verschillen in de onderwijsdoelen verduidelijken:

- ✓ Het algemeen secundair onderwijs (ASO). Hier wordt de nadruk gelegd op een ruime algemene vorming die vooral een stevige basis biedt voor het volgen van hoger onderwijs.
- ✓ Het technisch secundair onderwijs (TSO). De aandacht gaat vooral naar algemene en technisch-theoretische vakken. Na het TSO kan de jongere een beroep uitoefenen of verder studeren in het hoger onderwijs. Bij de opleiding horen ook praktijklessen.
- ✓ Het beroepssecundair onderwijs (BSO). Dit is een praktische onderwijsvorm waarin de jongere een specifiek beroep aanleert, terwijl hij ook algemene vorming ontvangt.

¹ Het kunst secundair onderwijs (KSO) zullen we omwille van de specificiteit en het beperkte aandeel in het totale secundaire onderwijs niet als een specifieke doelgroep opnemen.

Wij zijn van mening dat het, gelet op de eigenheid van de verschillende onderwijsvormen binnen het secundair onderwijs, aangewezen is voor elk van deze drie onderwijsvormen een representatieve studierichting als voorbeeld naar voren teschuiven. In paragraaf 4 van dit hoofdstuk geven we kort een overzicht van de meest specifieke kenmerken van de vooropgestelde studierichtingen. Eerst dringt zich evenwel nog volgende opmerking op bij de structuur van het Vlaamse secundair onderwijs.

In de eerste graad van het secundair onderwijs is er in principe geen sprake van een onderverdeling in onderwijsvormen. Het tweede leerjaar van de eerste graad bereidt de leerling evenwel voor om bij de overstap naar de tweede graad te kiezen voor de studierichting in één van de vier onderwijsvormen (algemeen, beroeps-, kunst- of technisch secundair onderwijs). Dit impliceert dat sommige studierichtingen als het ware (doch niet noodzakelijk) sterk aansluiten bij een latere onderwijsvorm. Bij het zoeken naar interessante casussen zullen wij, om een verdeling over graden en onderwijsvormen te realiseren, een deel van het tweede leerjaar van de eerste graad als kenmerkend beschouwen voor de hypothetische ASO-richting in de eerste graad van het secundaire onderwijs. Dit brengt ons bij volgende drie casussen: in het tweede leerjaar van de eerste graad A willen we ons specifiek toeleggen op de doelstellingen voor de vakken wiskunde en Nederlands; in de tweede graad van het technisch secundair onderwijs richten we ons op de studierichting Elektriciteit/Elektronica en in de derde graad van het beroepsonderwijs richten we ons op de studierichting Verzorging.

Voor deze drie studierichtingen zullen we een aantal doelstellingen naar voren schuiven waarvan we de evaluatie in het verdere verloop van dit onderzoeksrapport zullen illustreren.

3 Selectie van doelstellingen

Indien we bij het presenteren van concreet evaluatiemateriaal een selectie van doelstellingen als uitgangspunt nemen, dan moeten deze doelstellingen zorgvuldig worden gekozen. We willen daarom niet alleen een zo ruim mogelijke verdeling van de casussen tussen onderwijsvormen en graden realiseren maar ook een goede dekking van de doelen binnen elk van de casussen. Zonder uitvoerig in te willen gaan op de problematiek van doelstellingen, presenteren we hiertoe bondig een taxonomie van onderwijsdoelstellingen.

3.1 Waarom een taxonomie?

Het vooropstellen van doelstellingen kan leiden tot een zeer omvangrijke lijst zodat enige ordening noodzakelijk is. Zoals geldt voor de vele modellen die in de onderwijsliteratuur gehanteerd worden, is onderstaande taxonomie een hulpmiddel bij het analyseren van complexe onderwijsprocessen, hier in het bijzonder bij het inventariseren van onderwijsdoelen. Hierbij staan volgende twee functies voorop (Standaert & Troch, 1998):

- ✓ De taxonomie geeft leerkrachten een overzicht van de verschillende categorieën doelstellingen die moeten worden nagestreefd en geëvalueerd. Ze maakt het mogelijk na te gaan welke categorie doelstellingen een bepaalde evaluatiemethode evalueert en welke niet aan bod komen. Op die manier kan men achterhalen welke evaluatiemethoden gecombineerd moeten worden om het geheel te evalueren.
- ✓ De taxonomie maakt het mogelijk de samenhang te zien tussen de verschillende categorieën doelstellingen.
- ✓ Op dit punt willen we door gebruik te maken van een taxonomie voornamelijk verzekeren dat binnen de drie casussen alle relevante categorieën doelstellingen opgenomen worden.

3.2 Inhoud van de taxonomie

Het met diepgang beschrijven van mogelijke classificaties van doelstellingen valt uiteraard buiten het bestek van dit onderzoek. Toch kunnen we niet voorbij aan een korte toelichting bij de verschillende aspecten van de taxonomie. De taxonomie die wij zullen hanteren is gebaseerd op de classificatie van doelstellingen van Standaert & Troch (1998). Hetzelfde geldt voor de beschrijving van de inhoud ervan.

Centraal in de taxonomie staat het onderscheid tussen twee gedragsniveaus, namelijk kennis en vaardigheid. De organiserende vraag is dus eigenlijk welke handelingen men verwacht dat leerlingen met een bepaalde informatie kunnen uitvoeren.

3.2.1 Kennis

Kennis omvat de verworven informatie die op een georganiseerde manier in het geheugen is opgestapeld (Standaert & Troch, 1998). Hieronder wordt niet alleen het louter memoriseren maar ook het begrijpen van die informatie verstaan. Vandaar dat een onderscheid wordt gemaakt tussen geheugenkennis en inzichtelijke kennis.

- ✓ Bij **geheugenkennis** gaat het om het herkennen of zich herinneren van vooraf geleerde informatie. Deze informatie kan zowel betrekking hebben op feiten als op procedures.
- ✓ Leerlingen moeten informatie evenwel niet alleen herkennen of herinneren, ze moeten er ook betekenis aan kunnen geven. **Inzichtelijke kennis** heeft te maken met die bekwaamheid om uit een gegeven informatie de betekenis te achterhalen. Voorbeelden van activiteiten die getuigen van inzichtelijke kennis zijn: het omzetten van gegeven informatie, het interpreteren van gegeven informatie en het voorstellen van informatie.

3.2.2 Vaardigheden

Een vaardigheid is de bekwaamheid om op effectieve en efficiënte wijze iets te doen met de verworven informatie. Standaert & Troch (1998) maken hierbij een onderscheid tussen eenvoudige (nabootsende) en productieve vaardigheden enerzijds en tussen drie vaardigheidsgebieden (cognitief, motorisch en affectief) anderzijds. We zullen in de beschrijving van elk van de vaardigheidsgebieden dan ook telkens een onderscheid maken tussen twee vaardigheidsniveaus, namelijk het geautomatiseerd en het productief toepassen van iets.

1. Cognitieve vaardigheden

Leerlingen moeten het inzicht in een aangeleerde regel of procedure kunnen aantonen door deze aan te wenden in concrete en nieuwe taken en situaties. Leerlingen moeten een regel of procedure dus niet alleen kennen (cfr. inzichtelijke kennis) maar ook kunnen gebruiken in een nieuwe context. Men noemt dit het **geautomatiseerd toepassen van informatie**. Mogelijke voorbeelden hiervan zijn het correct toepassen van spellingsregels en het uitvoeren van het algoritme van een deling. Wanneer de procedure of het principe om een probleem(taak) op te lossen niet gegeven zijn, zullen de leerlingen deze zelf moeten vinden door te denken en te redeneren. Het lerende denken, of **productief toepassen van informatie**, is dan ook een belangrijke doelstelling van het onderwijs. Mogelijke voorbeelden hiervan zijn denkprocessen zoals analyserend denken, synthetiseren en beoordelen.

2. Motorische vaardigheden

Ook bij de motorische vaardigheden maken Standaert & Troch (1998) een onderscheid tussen twee toepassingsniveaus. Men spreekt van **geautomatiseerd bewegen** wanneer leerlingen aangeleerde bewegingen navolgen. De leerlingen kunnen hierbij terugvallen op reeds aangeleerde bewegingsactiviteiten zoals het opmaken van een ziekenbed of het gebruiken van een machine. Ook kennis en motorische vaardigheden kunnen in nieuwe situaties (met een probleemkarakter) worden toegepast. Dergelijk **productief bewegen** bestaat uit het combineren van denken en bewegen en uit zich in activiteiten zoals sporten, musiceren, dansen, handvaardigheid en dergelijke.

3. Affectieve vaardigheden

Net zoals bij de cognitieve en motorische vaardigheden kunnen we bij de affectieve vaardigheden een onderscheid maken tussen eenvoudig en meer productief affectief gedrag. Het onderscheid is evenwel niet zo eenvoudig en kan worden teruggebracht tot verschillende graden van verinnerlijking van affectief gedrag (Standaert & Troch, 1998). Hierbij is het onderscheid tussen de begrippen **rolgedrag** en **waardegedrag** van belang.

- Wanneer leerlingen hun gedrag schikken naar bepaalde gestelde normen, spreekt men van een **rolgedrag**. Men doet iets (niet) omdat het verwacht wordt in de groep waarvan

je deel uitmaakt. Een leerling kan bijvoorbeeld recht staan bij het binnenkomen van een leerkracht omdat er als hij dat niet doet een ongewenst gevolg dreigt (zoals een straf, boze blik of berisping). De leerling confirmeert zich aan verwachtingen die hij zelf niet zinvol vindt.

- Anders is het wanneer een gedrag wordt gesteld omwille van een gewilligheid die het gevolg is van een eigen beslissing. Als de leerling een gedrag stelt na een keuze op grond van een eigen waardesysteem spreekt men van **waardegedrag**. In vergelijking met een rolgedrag is er sprake van een sterkere vorm van verinnerlijking², zodat de leerling in kwestie als het ware onafhankelijk is van externe druk.

Een belangrijke categorie binnen onderwijsdoelstellingen betreft het vormen van affectieve vaardigheden in het algemeen en waardegedrag in het bijzonder. Het is in dit deel van affectieve doelstellingen dat we ook het ontwikkelen van attitudes moeten situeren. Attitudes moeten evenwel duidelijk onderscheiden worden van waarden. Standaert & Troch (1998, p. 68) schrijven dat een attitude een zekere mate van voorkeur of tegenzin is ten overstaan van een concrete situatie. Een waarde overstijgt die concreetheid en is als het ware een gewenste manier van gedragen. Uit één waarde kunnen dus vele attitudes voortvloeien. Bijvoorbeeld: de waarde objectiviteit resulteert in attitudes zoals 'tekorten en gebreken onderscheiden' en 'zich echt voordoen'.

De taxonomie met de verschillende categorieën doelstellingen die werden toegelicht, kunnen we schematisch als volgt voorstellen.

Figuur 2.2: Een taxonomie van onderwijsdoelstellingen

1. Kennis		a. Geheugenkennis Feitenkennis, procedurekennis	b. Inzichtelijke kennis Omzetten, interpreteren, voorspellen
2. Vaardigheden	Niveau	Routinematig toepassen	Probleemoplossend toepassen
	Cognitief	a. Een aangeleerde regel of procedure in een nieuwe situatie toepassen.	b. Productief toepassen van informatie: probleemoplossend denken, analyseren, synthetiseren, evalueren.
	Motorisch	c. Handelingen kunnen automatiseren.	d. Productief bewegen, combineren van denken en bewegen.
	Affectief	e. Rolgedrag: handelen op grond van normen.	f. Waardegedrag: handelen op grond van een eigen waardesysteem, attitudes.

4 Uitwerking van de casussen

Op basis van de hoger uitgewerkte taxonomie willen we voor elk van de vooropgestelde doelgroepen een selectie maken van doelstellingen waarvoor we doorheen het handboek voorbeeldmateriaal zullen ontwikkelen. Voor elke studierichting selecteren we een tweetal representatieve

² De graad van verinnerlijking kent een zestal stappen (gaande van rolgedrag naar waardegedrag): *aandacht opbrengen voor; zich schikken in; bereidheid tot reageren; waardering; organisatie van waarden; karaktereigenschap* (Standaert & Troch, 1998, p.58).

vakken. Daarnaast nemen we eveneens een aantal interessante vakoverschrijdende doelstellingen op. We baseren ons hiervoor indien mogelijk op eindtermen en bestaande leerplannen. Terloops willen we nogmaals benadrukken dat we met deze selectie van doelstellingen geen volledigheid nastreven. De doelstellingen worden gekozen met het oog op het vinden van voor leerkrachten leerrijke voorbeelden van evaluatie.

Voor elk van de casussen geven we eerst kort een toelichting over de betreffende studierichting. Vervolgens presenteren we telkens een selectie doelstellingen. Deze wordt gestructureerd op basis van de in de figuur 2.2 aangegeven onderverdeling.

4.1 Eerste graad – vakken Nederlands en Wiskunde

In het tweede leerjaar A van de eerste graadzullen we ons specifiek toeleggen op de doelstellingen voor de vakken wiskunde en Nederlands. Deze vakken oefenen de leerlingen vooral in het abstract denken, en dit zowel via de taal als via de wiskunde. Van leerlingen die deze studierichting succesvol doorlopen hebben, verwacht men dat ze volgende doelstellingen bereikt hebben. De structuur van de opsomming is gebaseerd op de verdeling binnen de taxonomie in figuur 2.2.

- ✓ Leerlingen kunnen het verband uitleggen tussen optellen en aftrekken, vermenigvuldigen en delen (1b).
- ✓ De leerlingen passen afspraken in verband met de volgorde van bewerkingen toe (2a).
- ✓ De leerlingen voeren de hoofdbewerkingen (optelling, aftrekking, vermenigvuldiging en deling) correct uit in de verzamelingen van de natuurlijke, de gehele en de rationale getallen (2a).
- ✓ De leerlingen kunnen vergelijkingen van de eerste graad met één onbekende oplossen (2a).
- ✓ De leerlingen kunnen vanuit tabellen met cijfergegevens het rekenkundig gemiddelde en de mediaan (voor niet-gegroepeerde gegevens) berekenen (2a) en hieruit relevante informatie afleiden (2b).
- ✓ De leerlingen kunnen een opgave herformuleren (2b).
- ✓ De leerlingen kunnen voor hun leeftijd bestemde formulieren invullen (2b).
- ✓ De leerlingen kunnen de volgende tekstsoorten produceren (2b): een informatief stuk voor bekende leeftijdgenoten, een verslag aan bekende volwassenen over een gegeven schoolse opdracht.
- ✓ De leerlingen kunnen luisteren naar een jeugdprogramma op radio en TV (2b).
- ✓ De leerlingen kunnen een goede schets of een aangepast schema van een wiskundig probleem maken (2d).
- ✓ De leerlingen beseffen dat in de wiskunde niet enkel het eindresultaat belangrijk is maar ook de manier waarmee het antwoord bekomen wordt (2f).
- ✓ De leerlingen ontwikkelen binnen de gepaste situaties een bereidheid om te reflecteren over hun eigen schrijven (taal, indeling, spelling) en om hun handschrift en lay-out te verzorgen (2f).

4.2 Tweede graad TSO - Elektriciteit/Elektronica

Deze studierichting biedt een polyvalente opleiding binnen de sector van de elektriciteit waarbij twee basisfunderingen worden gelegd; elektriciteit en elektronica. Het is een theoretisch-

technische opleiding. Dit wil zeggen dat men er enerzijds een stevige theoretische kennis opdoet van elektriciteit, elektronica, wiskunde en fysica. Anderzijds wordt een basis gelegd van praktische uitvoering. In de richting Elektriciteit/Elektronica leert de leerling gestructureerd inzichtelijk en creatief denken en handelen. Hij moet in staat zijn zelfstandig en met aandacht voor kwaliteit en milieu te werken. In de snel evoluerende hoogtechnologische leer- en werkomgeving moet de leerling vlot kunnen meegroeien. De optie elektriciteit/elektronica wil in de eerste plaats een voorbereiding aanbieden voor leerlingen die kiezen voor het hoger onderwijs. Afgestudeerden hebben evenwel ook verscheidene tewerkstellingsmogelijkheden.

De twee vakken waaruit we doelstellingen selecteren zijn 'praktijk elektriciteit' en 'electriciteit'. Daarnaast nemen we een aantal vakoverschrijdende doelstellingen op.

- ✓ Soldeertechnieken en -materialen kunnen toelichten (1a).
- ✓ De opbouw van een stof (moleculen, atomen, elektronen) kunnen toelichten (1a).
- ✓ Het dossier van een huisinstallatie (met verlichting, stopcontacten, aarding, schakelkast en dergelijke) kunnen toelichten (1b).
- ✓ Meettoestellen (Volt-, ampère- en Ohmmeter) kunnen gebruiken (2a).
- ✓ Fouten in een signalisatieschakeling (bvb. een bel) op een gestructureerd manier kunnen opzoeken (2b).
- ✓ De leerlingen kunnen een verslag maken van de metingen die zij zelf uitvoeren (2b).
- ✓ Elektriciteitsbuizen op maat brengen, ontbramen en buigen (2c).
- ✓ Fouten in een signalisatieschakeling (bvb. een bel) kunnen herstellen (2d).
- ✓ Spontaan de nodige veiligheidsvoorschriften toepassen (2f).
- ✓ Spontaan aandacht besteden aan de staat van de gebruikte gereedschappen (2f).

4.3 Derde graad BSO - Verzorging

De studierichting Verzorging bereidt leerlingen voor op een professionele loopbaan in de verzorgingssector. Centraal in deze studierichting staat de vorming tot geïntegreerd zorgverlener aan een zorgbehoevend persoon of kind, zowel op fysisch als op psycho-pedagogisch en sociaal vlak. Lessen en stages worden gecombineerd. De stages worden ingericht in kinderdagverblijven, kraamafdelingen, instellingen voor gehandicapten en bejaardentehuizen. In deze optie spelen dan ook de volgende componenten een zeer grote rol: een huishoudelijke en animatiecomponent (werken aan woon- en leefklimaat), een pedagogische, communicatieve component en een verzorgingstechnische component (verzorging).

Kwaliteitsvolle zorgverlening aan jonge en oudere zorgvragers vergt belangrijke attitudes, vaardigheden en inzichten. Volgende lijst doelstellingen is een selectie van leerplandoelen die vooropgesteld worden voor de vakken 'Omgangskunde' en 'Praktijk verzorging'. Daarnaast zullen we ons richten op de evaluatie van de stageactiviteiten van deze leerlingen.

- ✓ De leerlingen kunnen de soorten menselijke behoeften benoemen (1a).
- ✓ De leerlingen kunnen verschillende soorten verlies toelichten aan de hand van voorbeelden (1b).
- ✓ De leerlingen kunnen omgaan met zelfevaluatie bij het eigen verwerven van attitudes (2b).
- ✓ De leerlingen kunnen positieve en negatieve eigenschappen van de eigen persoon als verzorgende opnoemen (2b).

- ✓ De leerlingen kunnen de verzorgingsbehoeften in verband met hygiënische verzorging gebruiken (2c).
- ✓ De leerlingen kunnen verschillende aspecten van non-verbaal actief luisteren toepassen (2d).
- ✓ De leerlingen beheersen vaardigheden met het oog op creatief, animatief en expressief omgaan met baby's en peuters (2d).
- ✓ De leerlingen respecteren het beroepsgeheim tijdens de stagepraktijk (2f).
- ✓ De leerlingen leren omgaan met gedragsregels en gewoontes van andere culturen (2g).
- ✓ De leerlingen tonen begrip en gaan respectvol om met (de familie van) zorgvragers (2g).

Hoofdstuk 2: Kenmerken van goede evaluatie

1 Inleidend

Op het einde van dit deel willen we in staat zijn om zowel de kwaliteit van bepaalde evaluatievormen te beoordelen als om uit te maken in welke mate een evaluatievorm kan bestempeld worden als *alternatief*. Dit vereist dat we stilstaan bij noodzakelijke kenmerken van goede evaluatie in het algemeen en van alternatieve evaluatiemethoden in het bijzonder. Uit deze aanpak blijkt dat ook evaluatiemethoden die niet onder de noemer 'alternatieve evaluatievormen' vallen, kunnen bestempeld worden als goede evaluatie. Een uitvoerige bespreking van deze -niet alternatieve- evaluatievormen valt evenwel buiten het bestek van dit handboek. Dit neemt niet weg dat we bij het beschrijven en illustreren van de kenmerken van goede evaluatie en van het evaluatieproces de leerkracht concrete aandachtspunten zullen aanreiken die de kwaliteit van bestaande evaluatievormen -op een soms eenvoudige wijze- kunnen verbeteren.

Evaluatiemethoden en -instrumenten moeten voldoen aan een aantal algemene kwaliteitsvereisten vooraleer we de stap kunnen zetten naar het ontwikkelen van alternatieve evaluatievormen. Deze gedachtegang weerspiegelt zich in de verdere opbouw van dit deel. In hoofdstukken 2 en 3 bespreken we respectievelijk de eisen aan goede evaluatie en de vereisten om van alternatieve evaluatie te kunnen spreken. Het derde hoofdstuk behandelt met andere woorden de kenmerken die verschillende alternatieve evaluatiemethoden gemeenschappelijk hebben. Dit tweede hoofdstuk richt zich evenwel eerst op de eisen van goede evaluatie.

2 Eisen van goede evaluatie

Dat een goed bedoelde evaluatievorm niet altijd even succesvol is in het evalueren van de vooropgestelde doelen, blijkt op een leuke manier uit volgend krantenartikel (DS, 18 december 2001).

Analfabeet slaagt voor toelatingsexamens twee universiteiten

RIO DE JANEIRO - In Brazilië is een analfabeet geslaagd in de toelatingsexamens van twee verschillende privé-universiteiten. De 27-jarige bakker Severino da Silva slaagde voor een toelatingsexamen aan de faculteit der Letteren van een universiteit. De man, die momenteel in een parochie in de noordelijke buurt van de stad Rio de Janeiro leert lezen, klikte per toeval, zoals in een loterij, de juiste A, B, C, D en E aan in een multiple-choice proef. Vorige week raakte de bakker op ongeveer dezelfde manier toegelaten tot de rechtsfaculteit van een grote, eveneens private, universiteit van Rio. Zijn percentage goede antwoorden ontsloeg de man van een geschreven proef (De standaard on-line, 18 december 2001).

Gelet op de soms belangrijke beslissingen die worden gebaseerd op verzamelde evaluatiegegevens, is het noodzakelijk dat evaluatie aan een aantal kenmerken en kwaliteitseisen tegemoet komt. Deze eisen zijn van toepassing op alle vormen van evaluatie, gaande van een 'doorde-weekse' toets tot de zogenaamde alternatieve evaluatievormen. Met de aandacht voor goede evaluatie in het algemeen willen we benadrukken dat het tegemoetkomen aan een aantal eenvoudige eisen de kwaliteit van de huidige evaluatie sterk kan verbeteren. Ook traditionele evaluatievormen kunnen dus best kwaliteitsvol zijn -althans indien ze aan de eisen van goede evaluatie voldoen. Bovendien is het wellicht zo dat de stap naar zogenaamde alternatieve evaluatievormen voor leerkrachten die de kwaliteitseisen in dit hoofdstuk reeds zoveel mogelijk nastreven niet zo groot is. Of anders: dat leerkrachten die aandachtig tegemoet komen aan de eisen van goede evaluatie reeds onbewust (wat een aantal aspecten betreft) zogenaamd alternatief evalueren.

In algemene termen kunnen twee belangrijke elementen naar voren worden geschoven waaraan een goede evaluatie dient te voldoen, namelijk **doelmatigheid** en **billijkheid**. In volgende paragrafen worden deze beide eisen verder gespecificeerd en waar nodig geïllustreerd en gerelateerd aan de eigen evaluatie-ervaringen.

Terloops willen we benadrukken dat deze kenmerken van een goede evaluatie niet alleen van belang zijn bij het opstellen en uitvoeren van een evaluatie. Ook bij het beoordelen van de evaluatiegegevens zijn ze belangrijk. Niet alle eisen zijn evenwel steeds even relevant. Voornamelijk bij summatieve evaluaties spelen de kwaliteitseisen een rol. Een formatieve evaluatie hoeft immers niet steeds aan alle eisen van de doelmatigheid te voldoen (althans indien de resultaten niet in rekening gebracht worden bij een eindbeoordeling). Als leerkracht moet je dus het belang van de kwaliteitsvereisten bij elke evaluatie afwegen tegen de achtergrond van de doelstellingen en de context van de betreffende evaluatie. In de afrondende derde paragraaf van dit hoofdstuk staan we uitvoerig stil bij het toepassen van de doelmatigheids- en billijkheidseisen.

2.1 Doelmatigheid

Doelmatigheid kunnen we zeer kernachtig omschrijven: iets is doelmatig als het geschikt is voor het doel waarvoor het gemaakt is. Het verzamelen en beoordelen van evaluatiegegevens verloopt met andere woorden doelmatig indien beantwoord wordt aan het doel waarvoor de gegevensverzameling ontworpen werd. Hiertoe worden de volgende drie doelmatigheidseisen voorop gesteld: validiteit, betrouwbaarheid en efficiëntie³. Een doelmatige evaluatie beantwoordt met andere woorden aan kwaliteitseisen aangaande drie aspecten.

³ De beschrijving van deze kenmerken is gebaseerd op: Herman et al, 1992; Standaert & Troch, 1998; Janssen, 1999; Janssens et al., 2000; Sluyter, 1998; Gredler, 1999; Mabry, 1999; Moskal et al., 2000; McMillan, 2001.

2.1.1 Validiteit

De betekenis van validiteit als kwaliteitseis voor evaluatiemethoden willen we duiden aan de hand van volgende denkoefening.

Even stilstaan... 2.1: De validiteit van een evaluatie-instrument

Stel u even devolgende situatie voor. Een leerkracht omgangskunde is bezig met het ontwerpen van een eindexamen op basis waarvan beslist zal worden welke leerlingen al dan niet bekwaam zijn om door te stromen naar het 2^{de} leerjaar van de derde graad. Eén van de doelen van de toets is te achterhalen welke leerlingen reeds de juiste sociale vaardigheden verworven hebben om de opleiding verder succesvol te kunnen doorlopen. Actief luisteren is één van deze vaardigheden. In het evaluatie-instrument is er dan ook een vraag opgenomen die zich richt op actief luisteren. Deze luidt als volgt: *Welke zijn de vijf belangrijkste principes die aan de grondslag liggen van succesvol actief luisteren?*

Denkoefening: welke conclusies kan je als leerkracht vastkoppelen aan de score van leerlingen op deze vraag? Zijn de beslissingen die de leerkracht wil nemen op basis van de antwoorden van leerlingen op deze vraag volgens u geoorloofd?

De aanpak van de leerkracht in het voorbeeld schiet tekort omdat vaardigheden van leerlingen - in dit geval actief luisteren - worden beoordeeld aan de hand van weinig relevante evaluatiegegevens. Het gegeven dat een leerling de belangrijkste principes van actief luisteren (zoals het belang van de lichaamshouding, de gelaatsuitdrukking en bevestiging) kan neerschrijven, zegt weinig over het feit dat de leerling in kwestie de bedoelde vaardigheid ook effectief beheerst en gepast zal aanwenden. Daarom stelt men dat de validiteit van deze evaluatie ontoereikend is; er wordt niet geëvalueerd wat de leerkracht beoogt te evalueren.

De validiteit van een evaluatiemethode doelt op de eigenschap dat een evaluatiemethode al dan niet meet wat de gebruiker ervan bedoeld heeft ermee te meten. Anders geformuleerd: men moet de juiste conclusies kunnen trekken uit de evaluatieresultaten. Indien een evaluatie-instrument ontwikkeld en gebruikt wordt om bijvoorbeeld de mate van analytisch redeneren te meten, dan zou dit instrument ook de mate van analytisch redeneren moeten meten. Hoe meer het instrument dit doel benadert, hoe groter de validiteit. Validiteit is dus geen kwestie van alles of niets, maar wel van gradaties.

Alvorens het concept validiteit verder uit te diepen, moeten we de aandacht vestigen op volgende bedenking.

Even stilstaan... 2.2: Niet het instrument is al dan niet valide, wel de conclusies

In de evaluatieliteratuur spreekt men voortdurend over de validiteit van een evaluatiemethode of van een evaluatie-instrument. Ook in dit boek doen we dit. Eigenlijk is deze terminologie niet geheel correct. Een toets kan in strikte zin niet valide of invalide zijn. Validiteit is immers geen kenmerk van de toets zelf of algemener van het instrument en de procedure om evaluatie gege-

vens te verzamelen. Het is een kenmerk van de interpretaties die gemaakt worden op basis van de door de toets verkregen evaluatiegegevens en van het gebruik dat gemaakt wordt van de evaluatiegegevens (McMillan, 2001). Validiteit verwijst dus eigenlijk niet naar het evaluatie-instrument op zich, maar wel naar de accuraatheid van de gemaakte inferenties en van de beslissingen die worden gebaseerd op de evaluatiegegevens. Ondanks het feit dat dit een belangrijk nuance is, hanteren wij de gangbare woordenschat. Dit maakt de beschrijving van validiteit enigszins toegankelijker. Waar we schrijven dat een evaluatie niet valide is, bedoelen we met andere woorden dat de conclusies die men baseert op de evaluatiegegevens niet valide zijn.

Aangezien een evaluatie veel en uiteenlopende bedoelingen kan hebben, zijn er verscheidene validiteiten te onderscheiden, die een evaluatie in verschillende mate kan bezitten (terloops: in het licht van bovenstaande opmerking gaan we met deze uitspraak als het ware in de fout). In de literatuur komen dan ook veel verschillende benamingen en definities van validiteiten voor. Anderen stellen dan weer dat de verschillende validiteiten verschillende aspecten zijn van één en hetzelfde overkoepelend begrip.

Volgende denkoefening doet ons bij wijze van inleiding stilstaan bij de verschillende soorten van validiteit van een evaluatie-instrument.

Even stilstaan... 2.3: De verschillende soorten van validiteit van een evaluatie

Elk van deze vragen stelt een ander type validiteit aan de orde:

- ✓ Mag je op basis van evaluatiegegevens over een aantal doelstellingen leerlingen beoordelen over heel het curriculum? Waarom wel/niet?
- ✓ Kan men door leerlingen een aantal principes van defensief autorijden te laten opschrijven, evalueren of leerlingen de vaardigheid defensief autorijden beheersen? Waarom wel/niet?
- ✓ Kan je op basis van ervaringen tijdens praktijkoefeningen voorspellen of leerlingen ook spontaan bepaalde veiligheidsvoorschriften zullen naleven? Waarom wel/niet?
- ✓ Kan je op basis van lessen die onderwijzers in opleiding aan elkaar geven uitspraken doen over hoe de toekomstige onderwijzers zich in een echte klassituatie zouden gedragen. Waarom wel/niet?

Het antwoord op de vragen in bovenstaande denkoefening kan worden verduidelijkt door de verschillende vormen van validiteit te beschrijven. Om gestalte te geven aan didactische evaluatie zijn de volgende vier het meest relevant: inhoudsvaliditeit, begripsvaliditeit, predictieve validiteit en levensechtheid⁴.

a. Inhoudsvaliditeit

In onderwijssituaties wordt voornamelijk een belangrijke rol toegekend aan de inhoudsvaliditeit van een toets. Inhoudsvaliditeit van een evaluatiemethode is de eigenschap dat

⁴ Gebaseerd op: Thio, 1988; Standaert & Troch, 1998; Sluyter, 1998; Mc Tighe & Ferrara, 1996; en Mabry, 1999; Miller et al., 1998).

de opgaven een representatieve steekproef vormen van de te toetsen kennis of vaardigheden. Meestal doelt men hiermee op de representativiteit qua leerstofgebied. Men spreekt daarom ook van *doelstellingenrepresentativiteit*. Dit betekent dat de evaluatie niet verder mag reiken dan de behandelde leerstof terwijl ze deze toch ook voldoende moet dekken. Een goede evaluatie weerspiegelt dus de totale leerinhoud. In de praktijk worden vaak uitspraken gedaan over het beheersen van een bepaald vak op basis van een beperkt aantal vragen of een beperkte opdracht. Vandaar het belang van deze kwaliteitseis.

b. Begripsvaliditeit

Men spreekt van begripsvaliditeit als kan worden aangetoond dat de evaluatie het door de evaluator beoogde kenmerk van de leerling meet. Het evaluatie-instrument moet kunnen nagaan of de leerling de vooropgestelde kennis en vaardigheden bereikt heeft. De opgaven van de evaluatie moeten dan ook zo goed mogelijk overeenstemmen met de vooropgestelde doelstellingen. Men spreekt daarom ook van *doelstellingvaliditeit*. Bij wijze van voorbeeld: als men praktische vaardigheden wil evalueren, dan doet men dat niet aan de hand van een schriftelijke toets.

c. Predictieve validiteit

Predictieve validiteit is de eigenschap die een evaluatiemethode heeft als de evaluatiescore een criteriumvariabele kan voorspellen. Als met een toets bijvoorbeeld later studiesucces kan voorspeld worden, dan heeft die toets een bepaalde predictieve validiteit.

d. Levensechtheid

Tenslotte vermelden we kort dat ook de levensechtheid van een evaluatietaak soms als een maat voor validiteit wordt aangezien. De levensechtheid (of authenticiteit) van een evaluatie verwijst naar de eigenschap dat een evaluatieopdracht of -taak die de leerlingen/studenten moeten vervullen, op zich als betekenisvol en waardevol wordt ervaren.

De opsomming van de verschillende soorten van validiteit roept volgende bedenking op. Bij de voorgaande denkoefening en bij het doornemen van de verschillende validiteitsaspecten zijn we als het ware op een tegenstrijdigheid gestoten. We kunnen bovenstaande validiteitsproblematiek immers als een soort dilemma voorstellen. Men kan ervoor kiezen de inhoudsdomeinen, waarover men op basis van de evaluatievorm generaliserende uitspraken wil doen, te beperken. Op die manier kan men de begripsvaliditeit van de evaluatie verhogen. Doch: door de te evalueren domeinen te zeer te willen specificeren, zal men echter eveneens de validiteit van de evaluatie in haar geheel ondermijnen. De evaluatieve uitspraken die men erop kan baseren omtrent de totale set doelstellingen zullen immers aan inhoudsvaliditeit moeten inboeten. Op basis van deze bedenking komen we tot de conclusie dat een evaluatie moet balanceren tussen een hoge begripsvaliditeit en een hoge inhoudsvaliditeit. Het komt er dan ook op aan afhankelijk van de doelstellingen en de vooropgestelde functies van de evaluatie een optimale balans tussen begrips- en inhoudsvaliditeit te realiseren.

Om aan de validiteitseisen te voldoen stellen Standaert & Troch (1998) in het verlengde van de bovenstaande soorten validiteiten twee vuistregels voorop: 1. zorg ervoor dat de evaluatie werkelijk de kennis en vaardigheden meet zoals bedoeld en 2. maak dat de leerdoelen naar inhoud en taxonomisch niveau zo goed mogelijk zijn vertegenwoordigd in de gehele evaluatie. Een concrete uitwerking van deze vuistregels vindt u in de volgende figuur.

Figuur 2.3: De validiteit van een evaluatie bewaken

Beschouwen we bijvoorbeeld de doelstellingen in de studierichting BSO-Verzorging, dan kunnen beide vuistregels tot de volgende invulling leiden.

1. Verschillende doelstellingen in de casus hebben betrekking op probleemoplossende vaardigheden van leerlingen. Deze kunnen niet alleen worden beoordeeld aan de hand van een schriftelijke eindevaluatie (c.q. een schriftelijk examen). Om de vaardigheden van leerlingen te beoordelen doet men er dan ook best aan deze vaardigheden zelf in een zo authentiek mogelijke context te evalueren. Zo zijn een rollenspel of concrete stageactiviteiten veel beter geschikt om de vaardigheden actief luisteren en aandacht voor hygiëne te evalueren dan een schriftelijk examen. Met evaluatiegegevens over de kennis van procedures van hygiënische voorschriften kan men als leerkracht geen uitspraken doen over de toepassing van die procedures door de leerling als verzorgende. Om (begrips)valide uitspraken te doen komt het er dan ook op aan zorgvuldig na te gaan welke doelstellingen aan de hand van welke evaluatiemethode en -taak kunnen worden nagegaan. Op die manier kan je als leerkracht ervoor zorgen dat de evaluatie werkelijk de kennis en vaardigheden meet zoals bedoeld.
2. De verschillende (soorten) doelstellingen in de taxonomie moeten elk in de eigenlijke evaluatie vertegenwoordigd worden. Kennis van de soorten menselijke behoeften kan men eenvoudig schriftelijk toetsen. De vaardigheid om deze kennis te gebruiken vergt andere evaluatievormen. De mogelijkheid van leerlingen om voor zichzelf na te gaan of ze al dan niet in staat zijn die kennis te gebruiken weer andere. Elk van deze doelstellingen verdient aandacht bij het evalueren. Dit betekent dat men als leerkracht verscheidene bronnen van evaluatiegegevens zal moeten aanwenden. Uiteraard kunnen niet alle doelstellingen expliciet geëvalueerd worden. Een over-dachte en strategische keuze van de te evalueren doelstellingen en de hiervoor geschikte evaluatiemethoden is dan ook noodzakelijk. Op die manier kan je als leerkracht ervoor zorgen dat de verschillende leerdoelen zo goed mogelijk zijn vertegenwoordigd in de evaluatie.

2.1.2 Betrouwbaarheid

Om doelmatige conclusies te kunnen vastknopen aan evaluatiegegevens is naast validiteit eveneens betrouwbaarheid vereist. Ook al wordt validiteit beschouwd als de belangrijkste eis van een evaluatie, toch kan een evaluatie niet valide zijn zonder dat ze betrouwbaar is. Betrouwbaarheid is dus een noodzakelijke maar geen voldoende voorwaarde om van validiteit te kunnen spreken (Airasian, 1994; McMillan, 2001). Om het begrip betrouwbaarheid in te leiden, eerst even een kleine denkoefening.

Even stilstaan... 2.4: De betrouwbaarheid van een evaluatie-instrument

Iedereen is wellicht al eens in een situatie beland waarin men onder collega's duidelijk van mening verschilt over de prestatie van een bepaalde leerling. In dat geval geven verschillende beoordelaars met andere woorden uiteenlopende waarderingen aan éénzelfde leerlingprestatie. Ook één enkele beoordelaar kan echter ongewild tot verschillende beoordelingen van één prestatie komen. Zo kan er bij leerkrachten een aanzienlijk verschil bestaan tussen beoordelingen op verschillende momenten. Dergelijke elementen zijn aan de orde wanneer we spreken over betrouwbaarheid.

Denkoefening: Neemt u even een zoals boven beschreven concrete situatie in gedachten.

- ✓ Waaraan zijn de verschillen in beoordeling tussen verschillende leerkrachten toe te schrijven? En de verschillen tussen verschillende evaluatiemomenten?
- ✓ Welke gevolgen hebben deze problemen voor de kwaliteit van de evaluatie?
- ✓ Hoe kunnen dergelijke problemen worden voorkomen?

De betrouwbaarheid is de mate waarin men staat kan maken op de evaluatieresultaten⁵. Dat wil zeggen de mate waarin de scores consistent, nauwkeurig en reproduceerbaar zijn, kortom vrij van meetfouten. Net als validiteit is ook betrouwbaarheid dus in weze een kenmerk van de evaluatiescores en niet van de evaluatie of het instrument op zich. Toch zullen we ook hier spreken van betrouwbare evaluatie (zie de opmerking in Even stilstaan...2). Bij een betrouwbare evaluatie mag het evaluatieresultaat niet beïnvloed worden door storende factoren. Anders geformuleerd: een evaluatie-instrument is betrouwbaar als bij herhaald gebruik onder gelijke condities dezelfde resultaten worden gevonden. Indien een evaluatie-instrument bij gelijke condities niet tot hetzelfde resultaat komt, is de kans dus groot dat het evaluatieresultaat geen correcte weerspiegeling geeft van de te evalueren kennis en vaardigheden.

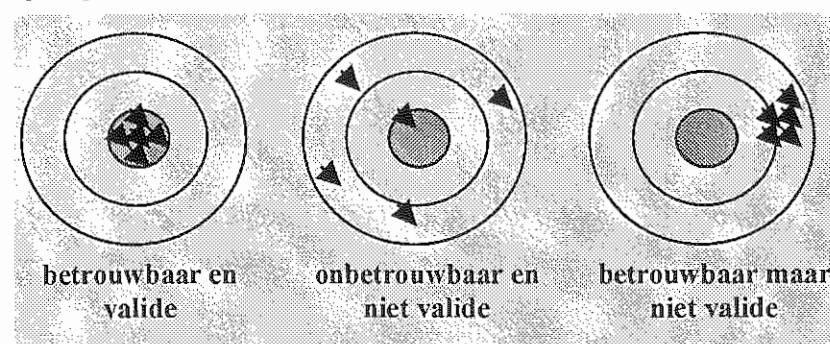
Figuur 2.4: Het verschil en het verband tussen betrouwbaarheid en validiteit

De concepten betrouwbaarheid en validiteit worden vaak met elkaar verward. Onthou daarom het volgende: betrouwbaarheid betreft enkel de vraag of de evaluatiescores consistent (eerlijk, correct, juist) zijn, validiteit houdt verband met de aard en de betekenis van de scores. Bij wijze van illustratie: de omtrek van iemands pols kan zeer betrouwbaar gemeten worden, maar het zou absurd zijn die omtrek te aanzien als een valide meting van de schrijfvaardigheden van die persoon.

Om het verband tussen validiteit en betrouwbaarheid te duiden doet McMillan (2001) een beroep op onderstaande figuren. Het centrum van de roos stelt de essentie voor van wat we willen evalueren (de te evalueren doelstelling). Stel nu dat we deze doelstelling vijfmaal evalueren en dat elke evaluatie resulteert in een pijltje. Als de verkregen evaluatiegegevens niet informeren over het vooropgestelde doel scoren ze niet in de roos en zijn ze niet valide. Wanneer de pijltjes dicht bij mekaar liggen wil dit zeggen dat steeds dezelfde beoordeling gegeven werd aan de evaluatie-

⁵ Gebaseerd op: Standaert & Troch, 1998; Mc Tighe & Ferrara, 1996; Janssens et al., 2000; Miller et al., 1998; McMillan, 2001.

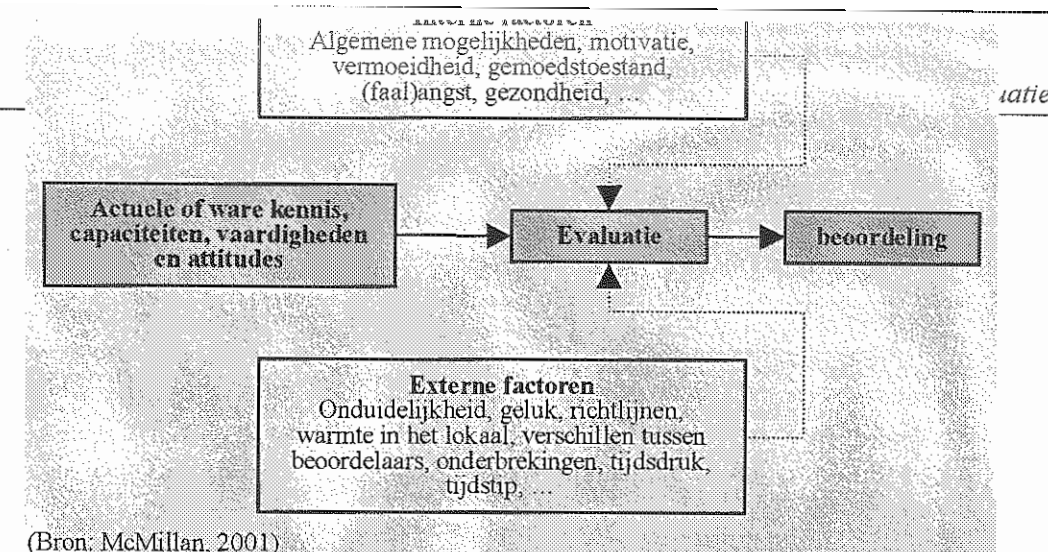
gegevens en dus dat de evaluaties betrouwbaar zijn. Op die manier wordt duidelijk dat een evaluatie tegelijkertijd betrouwbaar en niet valide kan zijn.



Betrouwbaarheid is belangrijk omdat de beslissingen die naar aanleiding van de evaluatie genomen worden, gebaseerd moeten zijn op betrouwbare gegevens die niet afhangen van allerlei toevalligheden. Interessant hierbij is de vraag naar de mogelijke oorzaken van een tekort aan betrouwbaarheid. Vaak liggen meetfouten of toevalligheden aan de basis van een gebrekkige betrouwbaarheid. Dergelijke storende factoren kunnen betrekking hebben op het evaluatie- en meetinstrument, op de meetsituatie, op de evaluatiekandidaat of op de beoordelaar. Volgende vragen plaatsen een aantal kritische bedenkingen bij de betrouwbaarheid van een bepaalde evaluatiemethode en maken duidelijk waar het bij betrouwbaarheid om draait.

- ✓ Wat het **evaluatie-instrument** betreft: zijn de vragen duidelijk geformuleerd zodat ze niet dubbelzinnig geïnterpreteerd kunnen worden? Zijn er voldoende vragen om de totale leerinhoud te evalueren en zijn de opgaven gespreid over diverse afnamemomenten zodat men een juiste weerspiegeling krijgt van de werkelijke kennis of vaardigheden van de leerling?
- ✓ Wat de **evaluatiesituatie** betreft: hebben het tijdstip en de duur van de evaluatie een invloed op de resultaten (maandagmorgen, vrijdagmiddag)? Hoe zit het met de tijdsdruk en de evaluatieomgeving (geluidsoverlast, concentratieproblemen)?
- ✓ Tenslotte, wat de **beoordelaar(s)** betreft: waarderen verschillende beoordelaars een bepaalde leerlingprestaties op een gelijkaardige manier, laten de beoordelaars niet relevante leerlinginformatie buiten beschouwing? Gaat men nauwkeurig en objectief tewerk of spelen subjectieve interpretaties en vooroordelen een rol?
- ✓ Wat de **evaluatiekandidaat** betreft: wat met een kort stresserend mondeling examen bij een faalangstige leerling, kan men dyslectische leerlingen op een betrouwbare manier schriftelijk evalueren? Zijn alle leerlingen wel even gemotiveerd om de evaluatieopdracht tot een goed einde te brengen?

Deze verschillende bronnen van meetfouten worden vaak onderverdeeld in twee groepen, namelijk interne factoren en externe factoren (zie fig. 2.5).



(Bron: McMillan, 2001)

Figuur 2.5: Bronnen van meetfouten bij het evalueren

Interne factoren situeren zich in de te evalueren leerling terwijl externe factoren betrekking hebben op invloeden buiten de leerling, zoals daar dus zijn het evaluatie-instrument, de evaluatiesituatie en de beoordelaars. Het is de combinatie van de verschillende groepen meetfouten die bepalend is voor de betrouwbaarheid van een evaluatiebeoordeling.

De mogelijkheid tot veralgemening is cruciaal voor evaluatiegegevens. Evaluatiescores worden dus pas relevant indien men er betrouwbare en valide veralgemeningen naar het bevraagde domein kan aan vastknopen. De score van een leerling op een bepaalde taak is in principe immers niet interessant op zich. De leerkracht Elektriciteit-Electronica is als het ware niet zozeer geïnteresseerd in het feit of een leerling een bepaalde taak succesvol kan afwerken. Hij wil op basis van de prestaties van de leerling op een specifieke taak uitmaken of de leerling in nieuwe, probleemstellende situaties in staat is hetzelfde te presteren. Om een grote veralgemeenbaarheid te kunnen realiseren, moeten doorgaans een aantal vormen van consistentie gerealiseerd worden. Schematisch ziet dit er als volgt uit.

Figuur 2.6: Verschillende aspecten van betrouwbaarheid

		Inhoudelijk (de vorm)	
		Hetzelfde	Verschillend
Tijdstip	Eén afname	Interne consistentie	Equivalentie
	Meerdere afnames	Stabiliteit	Stabiliteit en equivalentie

Zo is er sprake van consistentie op één enkel evaluatiemoment over parallelle taken en over taken die variëren qua inhoud en vorm maar toch hetzelfde inhoudsdomein evalueren. In het geval men de consistentie bekijkt over verschillende items die dezelfde inhoud nagaan en dit binnen één evaluatie spreekt men over *interne consistentie*. Wanneer twee inhoudelijk verschillende evaluaties gegeven worden op hetzelfde ogenblik kan men de equivalentie of *degelijkheid* van de beoordelingen nagaan. Ook de *stabiliteit* van prestaties op dezelfde taak over verschillende momenten is in dit opzicht belangrijk. Het gebruik van verschillende evaluatievormen op verschillende tijdstippen kan dan weer zowel informeren over de gelijkwaardigheid als de stabiliteit.

Kortom: niet alleen een voldoende consistentie tussen verschillende beoordelaars is belangrijk, ook consistentie over verschillende tijdstippen en over verschillende taken die hetzelfde beogen te meten zijn belangrijke aspecten van betrouwbaarheid. In dit laatste geval spreekt men ook wel eens van transfer en veralgemeenbaarheid (Herman et al., 1992). Birenbaum (1995) merkt hierbij op dat de betrouwbaarheid over taken in realiteit vaak minder hoog blijkt dan de betrouwbaarheid over beoordelaars. Of dit al dan niet een probleem is, hangt af van de doelen die men beoogt met de betreffende evaluatievorm. De eisen die men aan de betrouwbaarheid van een evaluatie stelt, worden immers gedetermineerd door de beslissingen en veralgemeningen die men op de evaluatiegegevens wil baseren (zie paragraaf 3).

2.1.3 Efficiëntie

Naast validiteit en betrouwbaarheid is de doelmatigheid van een evaluatie eveneens afhankelijk van de efficiëntie. Het gaat hierbij voornamelijk om een aantal praktische eisen zoals de tijds- en kosteninvestering (Standaert & Troch, 1998; Herman et al., 1992). De praktische mogelijkheden van leerkrachten om evaluatietaken te ontwerpen en de beschikbare tijd die aan de eigenlijke evaluatie kan besteed worden, zijn immers niet oneindig. Ook de vragen of de betreffende evaluatiemethode wel de meest geschikte is om de vooropgestelde doelstellingen te evalueren en of de evaluatiemethode wel degelijk de noodzakelijke informatie oplevert, zijn hier aan de orde.

2.1.4 Doelmatigheid: bemerkingen en criteria

We hebben reeds benadrukt dat leerkrachten bij het evalueren moeten streven naar een goede balans tussen verschillende soorten validiteit. Hetzelfde geldt eigenlijk voor de drie aspecten van doelmatigheid onderling (validiteit, betrouwbaarheid en efficiëntie). Het realiseren van uiterst valide en betrouwbare evaluaties vergt een intensieve tijds- en kosteninvestering van de leerkracht. Het komt er dan ook op aan als leerkracht een zo groot mogelijke efficiëntie na te streven. Dit kan betekenen dat de evaluatie binnen aanvaardbare grenzen aan validiteit en betrouwbaarheid moet inboeten. Maar uiteraard ook andersom: wanneer de evaluatie een zeer hoge doelmatigheid (validiteit, en betrouwbaarheid) vereist, kan je als leerkracht niet voorbij aan een zekere tijdsinvestering.

Aan het einde van de beschrijving van de doelmatigheidseisen om van een goede evaluatie te kunnen spreken, brengen we een aantal essentiële criteria samen. Die kunnen leerkrachten helpen uitmaken of een evaluatie al dan niet tegemoet komt aan deze belangrijke kwaliteitseisen. De lijst is niet uitputtend, maar wil een selectie van vragen presenteren die men zich kan stellen om de doelmatigheid van een evaluatie kritisch te bekijken.

Figuur 2.7: Kritische vragen om de doelmatigheid van een evaluatie te beoordelen

- ✓ Brengt de evaluatie de meest relevante gegevens over de te waarden leerlingenprestaties aan het licht?
- ✓ Stimuleert de evaluatie een ruime blik op het curriculum? Omvat de evaluatie een breed scala van inhoud en vaardigheden?
- ✓ Is er voldoende tijd voorzien zodat de evaluatie de bedoelde capaciteiten van leerlingen meet in plaats van hun vaardigheden om toetsen af te leggen. Voorziet de evaluatie in taken waarvan de oplossing niet vooraf kan worden ingestudeerd?
- ✓ Richt de evaluatie zich op de sleutel doelstellingen van het betreffende vak? Worden tijdens de evaluatie alle mogelijke irrelevante problemen vermeden (zoals taalgebruik en complexe zinsbouw)?
- ✓ Hebben de leerlingen de noodzakelijke achtergrondinformatie om de evaluatietaken op te lossen?
- ✓ Biedt de evaluatie informatie over de vraag of de evaluatieresultaten consistent zijn over beoordelaars en evaluatiemomenten?
- ✓ Is de evaluatie (en het scoren ervan) de geïnvesteerde tijd waard? Is ze administratief en qua kosten haalbaar?

2.2 Billijkheid

Naast doelmatig moet een evaluatie ook billijk zijn. De verschillende deelnemers aan een evaluatie moeten dezelfde behandeling krijgen. Dit is in een notendop wat met de eis van billijkheid wordt bedoeld (Herman et al., 1992; Standaert & Troch, 1998; McMillan, 2001). Je kan spreken

van een rechtvaardige evaluatie indien er sprake is van objectiviteit, doorzichtigheid en normering. Deze drie billijkheidsvereisten lichten we kort toe.

2.2.1 Objectiviteit

In een objectieve evaluatie hebben leerlingen de kans om hun prestaties te demonstreren zonder dat niet relevante individuele kenmerken van invloed zullen zijn op de beoordeling. Elke leerling moet met andere woorden dezelfde kansen krijgen. Volgende voorbeelden van evaluatieopdrachten verduidelijken het belang van objectiviteit.

Even stilstaan... 2.5: Aspecten van objectiviteit

Leerlingen worden gevraagd een kortverhaal te schrijven over een jongen die hard oefent om een goed voetballer te worden zodat hij in het schoolteam kan aantreden. Denk je dat meisjes dezelfde kansen hebben om hier een leuk verhaal rond te verzinnen als jongens?

Leerlingen beantwoorden een aantal vragen ter evaluatie van comprehensief lezen. De tekst handelt over de historische ontwikkeling van de onfeilbaarheid van de paus. Denk je dat katholieke leerlingen dezelfde kansen zullen hebben om tot goede antwoorden te komen als leerlingen met een ander geloof of vrijzinnige leerlingen?

Een leerkracht Scheikunde beoordeelt de prestaties van leerlingen normgerefeerd. Dit betekent dat hij de prestaties van de individuele leerlingen afzet tegen de gemiddelde prestaties. Wie beter scoort dan gemiddeld krijgt een goede score, wie lager scoort een slechte. Denk je dat de minder sterke leerlingen dezelfde kansen zullen hebben om een goede beoordeling te behalen? Betekent dit dan dat hun prestaties altijd ondermaats zijn?

Uit de denkoefening blijkt dat objectiviteit verschillende aspecten heeft. Deze vereisten aan de objectiviteit van een evaluatie sluiten zeer nauw aan bij de aandachtspunten die aan bod kwamen bij betrouwbaarheid. De invalshoek is evenwel anders. Met objectiviteit wordt benadrukt dat elke leerling recht heeft op een rechtvaardige en gelijke evaluatie. Met betrouwbaarheid wil men nastreven dat de evaluatiegegevens naar een ruimer domein kunnen veralgemeend worden. Volgende alinea verduidelijkt kort de belangrijkste aspecten van objectiviteit.

Onder objectiviteit verstaan we in de eerste plaats dat de beoordelaar geen enkele invloed mag hebben op de evaluatieresultaten van de geëvalueerde. Het mag met andere woorden voor de resultaten geen verschil maken wie instaat voor de afname, correctie en waardering van de evaluatie. Naast de invloed van beoordelaar moet ook de eventuele invloed van andere – niet aan de te evalueren prestaties gerelateerde – variabelen zoveel mogelijk teniet worden gedaan. Het is immers niet uitgesloten dat verschillende leerlingen éézelfde taak omwille van bepaalde persoonlijke kenmerken (zoals motivatie of geslacht en sociaal-economische achtergrond) anders aanpakken. Zo zou het herhaaldelijk gebruik van sportieve prestaties in rekenvraagstukken bepaalde leerlingen - die niet in sport geïnteresseerd zijn - kunnen irriteren. En actueler is de vraag of verschillende leerlingen wel in gelijke mate toegang hebben tot informatie- en communica-

tietechnologie (het internet) wanneer ze bijvoorbeeld de opdracht krijgen een spreekbeurt voor te bereiden. Als leerkracht moet men steeds waakzaam zijn opdat dergelijke verschillen tussen leerlingen niet leiden tot vertekeningen van de leerlingenprestaties.

2.2.2 Doorzichtigheid

Een evaluatiemethode wordt gekenmerkt door doorzichtigheid (of transparantie) indien de leerling alle informatie krijgt die hij nodig heeft om te zorgen voor een optimale voorbereiding op en een adequate uitvoering van de evaluatieopdracht (Standaert & Troch, 1998). Dit impliceert dat er duidelijkheid moet geschapen worden omtrent de vooropgestelde doelstellingen (wat geëvalueerd wordt) en over de criteria en standaarden die zullen gehanteerd worden om de prestaties te evalueren (hoe geëvalueerd wordt). De leerlingen moeten tevens een goed zicht hebben op de manier waarop een bepaald resultaat tot stand komt. Hetzelfde geldt voor de eventuele gevolgen van de evaluatie. De verwachtingen over wat, en op welke manier de leerling zal geëvalueerd worden, zijn immers in sterke mate bepalend voor de manier waarop men zich op de evaluatie voorbereidt.

De leerkracht kan op verschillende manieren de doorzichtigheid van een evaluatie garanderen. Te denken valt aan het geven van voorbeeldvragen, aan het expliciet benadrukken van de essentiële aspecten van de leerinhoud, aan het benadrukken van sleutelvaardigheden en aan het creëren van een sfeer van vertrouwen waarin informatie gegeven wordt over de aard van de evaluatie en de soort vragen en opdrachten. Dit is evenwel niet vanzelfsprekend. Het risico bestaat immers dat leerlingen zich bij een te grote doorzichtigheid beperken tot het louter studeren in functie van voorbeeldvragen en de benadrukte kern van de inhoud (Janssens et al., 2000). Op die manier dreigt de evaluatie opnieuw aan validiteit te moeten inleveren.

2.2.3 Normering

De derde voorwaarde om de billijkheid van een evaluatie te garanderen, is normering. Met deze voorwaarde wordt benadrukt dat men de evaluatieve beslissingen inhoudelijk moet kunnen rechtvaardigen. Vooral het gebruik van zogenaamde evaluatiestandaarden kan hiertoe bijdragen. Alle betrokkenen moeten weten welke prestaties verwacht worden om een bepaalde beoordeling te krijgen. Concreter: wanneer is een verslag dat leerlingen Elektriciteit/Elektronica maken van de metingen die zij uitvoerden voldoende goed om te concluderen dat de doelstelling op voldoende wijze beheerst wordt? Welke prestatie leidt tot een uitmuntende score? Dergelijke duidelijkheid over de evaluatiestandaarden maakt het mogelijk de evaluatieve beslissingen te stoffen en te rechtvaardigen. Op het belang van en het ontwerpen van evaluatiestandaarden gaan we dieper in bij het bespreken van het evaluatieproces (zie deel 3).

2.2.4 Billijkheid: opmerkingen en criteria

Ook voor de billijkheidseisen kunnen we in de vorm van kritische vragen een aantal criteria formuleren. Deze kunnen helpen uitmaken of een evaluatie al dan niet tegemoet komt aan de vooropgestelde kwaliteitseisen. De lijst is niet uitputtend maar wil een selectie van vragen presenteren die men zich kan stellen om de billijkheid van een evaluatie kritisch te bekijken.

Figuur 2.8: Kritische vragen om de billijkheid van een evaluatie te beoordelen

- ✓ Is de evaluatie niet beïnvloed door culturele, sociaal-economische en geslachtsgebonden stereotypen?
- ✓ Hebben de leerlingen vooraf een duidelijk zicht op de onderwijsdoelen en de aard van de evaluatie?
- ✓ Hadden de leerlingen effectief de kans om zich het geëvalueerde eigen te maken, om te oefenen? Richt de evaluatie zich op kennis en vaardigheden waartoe leerlingen effectief de kans hadden om ze doorheen de instructie aan te leren?
- ✓ Hebben de leerlingen de nodige vaardigheden om de evaluatie tot een goed einde te brengen? Hebben ze inhoudelijk de nodige voorkennis?
- ✓ Voorziet de evaluatie in expliciete criteria voor het scoren van prestaties van leerlingen? Zijn deze criteria voor alle betrokkenen doorzichtig en beschikbaar?
- ✓ Kunnen de beoordelaars de beoordeling die ze geven inhoudelijk rechtvaardigen?

Net zoals bij de doelmatigheidseisen komt het er bij de billijkheidseisen op aan een zekere realiteitszin in acht te nemen en om afhankelijk van de doelen/functie van de evaluatie meer of minder belang te hechten aan de verschillende elementen. In de volgende paragraaf staan we stil bij het differentiële belang van zowel de doelmatigheids- en billijkheidscriteria en gaan we na hoe je als leerkracht deze kwaliteitseisen kan bewaken.

3 De evaluatie geëvalueerd

Wanneer men de evaluatie op zich tot onderwerp maakt van evaluatie spreekt men van meta-evaluatie (Scriven, 1969). De achterliggende idee stelt dat meta-evaluatie kan gebeuren aan de hand van een logica die gelijkaardig is aan die van het evalueren van leerlingen. Dit betekent bijvoorbeeld dat gelijkaardige methoden en instrumenten kunnen gebruikt worden en dat dezelfde kwaliteitseisen van kracht zijn. Op dit punt verstaan we onder meta-evaluatie dat op een systematische manier wordt nagegaan of de evaluatie effectief aan de noodzakelijke kwaliteitseisen voldoet. Over deze kwaliteitseisen dienen nog twee vragen te worden gesteld: 1. hoe strikt moeten ze worden toegepast en 2. hoe kan de kwaliteit ervan bewaakt worden? Inzicht in deze opmerkingen bij de kwaliteitseisen is verhelderend om op een realistische en correcte wijze te kijken naar de eigen evaluatiepraktijk.

3.1 Toepassing van doelmatigheids- en billijkheidscriteria

De hoger geformuleerde kwaliteitseisen zijn van belang in elke fase van het evaluatieproces, d.w.z. bij het verzamelen, beoordelen en rapporteren van evaluatiegegevens. Een evaluatie kan dus maar kwaliteitsvol zijn indien de doelmatigheids- en billijkheidscriteria doorheen het ganse proces gerealiseerd worden. Toch gaat het hierbij niet om een kwestie van alles of niets. Het is raadzaam de kwaliteitseisen te benaderen als gradaties van meer of minder validiteit, betrouwbaarheid, doorzichtigheid terwijl desalniettemin getracht wordt zoveel mogelijk aan alle eisen tegemoet te komen. In principe is het dus aangewezen om steeds een optimale kwaliteit na te streven, goed wetende dat ook de efficiëntie van een evaluatie moet bewaakt worden. Een aantal redenen maken het echter noodzakelijk om toch op een flexibele manier met de doelmatigheids- en billijkheidseisen om te springen.

1. Flexibiliteit omwille van de verscheidene functies van evaluatie

Bij het beschrijven van de verschillende aspecten van doelmatigheid en billijkheid werd reeds kort stilgestaan bij de flexibiliteit waarmee bovenstaande evaluatievereisten toegepast dienen te worden. Een evaluatie die tot doel heeft te beslissen of leerlingen al dan niet geslaagd zijn voor een bepaald vak en waarvan de resultaten worden meegenomen bij het attesteren van de betreffende leerling, moet veel strikter aan deze eisen voldoen dan een begeleidend evaluatiemoment dat tot doel heeft tussentijdse feedback te genereren. Het spreekt voor zich dat de eventuele negatieve gevolgen van het niet volledig tegemoetkomen aan de evaluatie-eisen in het laatste geval veel minder verregaand zullen zijn. De doelen van de evaluatie zijn dus bepalend voor de striktheid waarmee de beschreven eisen dienen te worden toegepast. Hierbij moet worden opgemerkt dat voorgaande bemerking voornamelijk geldt voor de betrouwbaarheid van evaluaties. De (inhouds)validiteit blijft immers steeds van belang omwille van de leerdoelgerichte feedback en begeleiding die men wil baseren op de formatieve evaluatie.

Uit deze reflectie onthouden we voornamelijk dat de verschillende kwaliteitseisen naargelang de evaluatiecontext flexibel moeten worden aangewend.

2. Flexibiliteit omwille van de spanning tussen kwaliteitseisen

Bij het presenteren van de verschillende soorten validiteit zijn we op een tegenstrijdigheid gestoten die we als een soort dilemma hebben voorgesteld: door het beperken van de inhoudsdoelstellingen waarover men uitspraken wil doen, kan men bijvoorbeeld de begripsvaliditeit van de evaluatie verhogen. Daartegenover staat dat door de te evalueren domeinen te zeer te willen specificeren uitgerekend aan inhoudsvaliditeit wordt ingeboet. Een evaluatie moet met andere woorden balanceren tussen een hoge begripsvaliditeit en een hoge inhoudsvaliditeit, wat dus een zekere flexibiliteit impliceert. Dit flexibiliteitscriterium situeert zich *binnen* de validiteit als kwaliteitseis. Ook *tussen* een aantal kwaliteitseisen stelt zich een gelijkaardig probleem. Bijvoorbeeld: een hoge validiteit gaat vaak ten koste van de betrouwbaarheid en de objectiviteit van de evaluatie. En ook andersom: wil men een hoge mate van betrouwbaarheid of een evaluatie die zoveel mogelijk vrij is van meetfouten, dan bestaat het gevaar te vervallen in het meten van reproductieve kennis. Dergelijke kennis kan dan wel op een eenvoudige en eenduidige ma-

nier beoordeeld worden maar zulks doet anderzijds dan weer afbreuk aan de validiteit van de evaluatie. Het komt er dus op aan een geschikte balans te vinden tussen de verschillende eisen. Ook hierin speelt de functie van de op te zetten evaluatie een cruciale rol.

3. Flexibiliteit omwille van de eisen zelf

We stelden reeds dat de algemene kwaliteitseisen moeten opgevat worden als continua. Een concrete evaluatie kan in meer of mindere mate aan de verschillende eisen tegemoet komen. Het gaat hierbij steeds om een zo goed mogelijk tegemoetkomen aan, aangezien het in principe zinloos is om te spreken van een perfecte doelmatigheid of een perfecte billijkheid. Het is immers onmogelijk om een perfect betrouwbare en een perfect valide evaluatie op poten te zetten. In beperkte mate zullen er steeds wel meetfouten en tekortkomingen zijn. Het doel van een evaluatie moet er daarom steeds in bestaan de evaluatie zo doelmatig mogelijk te doen verlopen. Dit kan bijvoorbeeld eveneens inhouden dat men op zoek gaat naar een doelmatigheidsniveau dat als minimum vooropgesteld wordt.

De boodschap die we willen overbrengen stelt dat bij elke evaluatie een zo hoog mogelijk kwaliteitsniveau wordt nagestreefd. Toch is het belangrijk bij het toepassen van de doelmatigheids- en billijkheidseisen de nood aan een zekere flexibiliteit in het achterhoofd te houden. In de eerste plaats omdat het in wezen onmogelijk is op een perfecte manier aan al de kwaliteitseisen te voldoen. Bovendien impliceert het tegemoetkomen aan de ene kwaliteitseis dat bij een andere aan kwaliteit moet worden ingeboet. Tenslotte is het niet in alle evaluatiecontexten even belangrijk om in dezelfde mate aan alle eisen tegemoet te komen.

Maar om af te ronden nogmaals: indien de efficiëntie van de evaluatie er niet door in het gedrang komt, is het uiteraard aangewezen de kwaliteit zo hoog mogelijk te maken. In de volgende paragrafen zetten we een aantal werkwijzen op een rij die kunnen bijdragen tot het bewaken en optimaliseren van de verschillende kwaliteitseisen.

3.2 Hoe de kwaliteit van een evaluatie bewaken?

Hoe kan je als leerkracht de kwaliteit van de gebruikte evaluaties bewaken? Precies door deze evaluaties zelf op regelmatige basis te onderwerpen aan een evaluatie. Bij het beschrijven van de verschillende kwaliteitseisen hebben we een aantal kritische vragen gepresenteerd die men zich als leerkracht kan stellen. Hieronder stellen we een aantal concrete werkwijzen en voorwaarden voor om deze kwaliteitseisen te garanderen. We maken hierbij een onderscheid tussen initiatieven die je als individuele leerkracht kan nemen en initiatieven waarbij de inbreng van derden vereist is.

3.2.1 De kwaliteit bewaken in je eentje

Zowel de validiteit als de betrouwbaarheid van didactische evaluaties kunnen door de individuele leerkracht op een vrij eenvoudige manier bewaakt worden. Hetzelfde geldt overigens voor de billijkheid van een evaluatie.

1. De validiteit bewaken

Wanneer leerlingen op grote schaal geëvalueerd worden, dan bestaan er procedures om aan de hand van correlatiecoëfficiënten uitspraken te doen over de validiteit van een evaluatie-instrument. Daar heb je als individuele leerkracht echter weinig aan. Toch kunnen verschillende suggesties naar voor worden geschoven om de validiteit van een evaluatie te bewaken. Zo kan je nagaan of het op verschillende manieren evalueren van dezelfde doelstellingen toch tot dezelfde resultaten leidt (vorm van de zogenaamde split-halfmethode). Of je kan de scores van de leerlingen waarvan je verwachtte dat ze hoge scores zouden behalen vergelijken met de scores van de groep leerlingen waarvan je verwachtte dat ze laag zouden scoren. Je kan ook de scores van de leerlingen vóór de instructie met deze van na de instructie vergelijken (McMillan, 2001).

2. De betrouwbaarheid bewaken

De betrouwbaarheid van een evaluatie wordt doorgaans bepaald door de grootte te schatten van de meetfout die de evaluatiescore kenmerkt. Hoe kleiner deze meetfout, des te betrouwbaarder de evaluatiebeoordeling. Onthou hierbij wel dat elke beoordeling gekenmerkt is door een zekere foutenmarge. Als individuele leerkracht kan je desalniettemin verscheidene richtlijnen en aandachtspunten in acht nemen om meetfouten te minimaliseren. We denken hierbij bijvoorbeeld aan volgende elementen.

- De ideale manier om de betrouwbaarheid van een evaluatie te beoordelen bestaat erin een evaluatie te organiseren en later hetzelfde te evalueren met een ander instrument. Op die manier kan de stabiliteit en de equivalentie van beoordelingen worden nagegaan. Het is echter weinig realistisch van leerkrachten te verwachten om een evaluatie tweemaal te organiseren simpelweg om de betrouwbaarheid van de evaluatie te beoordelen. Toch is het zinvol dit als ideaal in het achterhoofd te houden (McMillan, 2001). Leerkrachten kunnen over een bepaalde tijdperiode immers verscheidene evaluaties geven die elk op zich misschien weinig betrouwbare evaluatiegegevens opleveren, terwijl het totaal van de evaluaties wel een algemene en betrouwbare beoordeling mogelijk maakt. Wanneer leerlingprestaties niet consistent zijn over verschillende evaluatietaken en dito evaluatiemomenten, dan moet dit de leerkracht extra aandachtig maken voor het effect van eventuele meetfouten. Het gebruik van verschillende evaluatietaken met voldoende items en dit op verschillende tijdstippen is dus een goede strategie om de betrouwbaarheid van je evaluatiebeoordeling te bewaken.
- Verder is het gebruik van evaluatiecriteria en rubrieken zeer doelmatig (zie deel 3). Belangrijk hierbij is deze zowel voor de leerlingen als voor de verschillende beoordelaars beschikbaar te maken. Om een betrouwbare evaluatie te garanderen komt het er op aan

zorgvuldig het evaluatieproces, zoals we dat in volgend deel zullen beschrijven, te doorlopen. Dit betekent bijvoorbeeld ook dat men duidelijkheid schept over de onderwijsdoelen en dat men de objectiviteit bij het beoordelen tracht te optimaliseren.

- Leerkrachten zijn zich vaak bewust van het feit dat bepaalde evaluaties voor bepaalde leerlingen geen betrouwbare data opleveren. Wanneer een leerling afgeleid is van een evaluatieopdracht omwille van familiale of gezondheidsproblemen dan weet je als leerkracht dat een normale beoordeling geen betrouwbare gegevens over die leerling oplevert. Een zekere openheid, en een veilige en flexibele leer- en evaluatieomgeving zijn noodzakelijk om dergelijke meetfouten aan het licht te brengen. De invloed van sommige interne beïnvloedingsfactoren kan je als leerkracht enigszins controleren. Zo kan je leerlingen motiveren om hun best te doen en de evaluaties differentiëren bij bijvoorbeeld faalangstige en dyslectische leerlingen.

3. De billijkheid bewaken

Als individuele leerkracht kan je de billijkheid, c.q. objectiviteit, van een evaluatiebeoordeling bewaken door jezelf de volgende vraag te stellen: presteren alle leerlingen met hetzelfde niveau qua kennis, inzicht en vaardigheden wat de te evalueren doelen betreft gelijkaardig en dit ongeacht hun onderlinge verschillen op andere gebieden. Als het antwoord 'ja' is, dan zal de evaluatie waarschijnlijk objectief zijn; indien het antwoord 'neen' luidt, is de evaluatie niet objectief. In dit laatste geval spelen niet-relevante leerlingenkenmerken zoals geslacht, voorkennis, sociaal-economische achtergrond, motivatie en dergelijke een ongeoorloofde rol in de beoordeling.

3.2.2 Hoe anderen betrekken in de kwaliteitsbewaking

Als leerkracht secundair onderwijs kan je verschillende groepen derden aanspreken om gezamenlijk de kwaliteit van de gebruikte evaluaties te verbeteren. We denken hierbij in de eerste plaats aan collega's en leerlingen, maar ook aan externen. We geven voor elke groep kort aan welke rol ze kunnen spelen.

1. De rol van externen

De rol van externen ten aanzien van de kwaliteitsbewaking kan zowel indirect als direct worden ingevuld.

Indirect in die zin dat externen leerkrachten op de hoogte kunnen houden van relevante evoluties. Op basis van die informatie kunnen leerkrachten zelf de doelmatigheid van hun evaluaties bewaken. Om aan de voorgestelde kwaliteitseisen te voldoen, is het immers belangrijk dat leerkrachten aandacht hebben voor ontwikkelingen binnen de toekomstige professie van hun leerlingen. Met professie bedoelen we zowel beroepsspecifieke vaardigheden als vaardigheden die noodzakelijk zijn om verdere studies aan te vatten. Afhankelijk van de studierichting waarin ze werkzaam zijn, moeten leerkrachten zicht hebben op wat in beide van essentieel belang is. Op die manier kan de validiteit van een evaluatie bewaakt en geoptimaliseerd worden. Dit uit zich in de vraag of de evaluatie wel relevant geachte inhouden en vaardigheden centraal stelt. Bo-

vendien informeert een regelmatige consultatie van de toekomstige werkvloer en studievloer ook over de standaarden die gehanteerd worden, in die zin dat je als leerkracht een beeld krijgt van wat als een voldoende goede prestatie vooropgesteld wordt.

Externen kunnen echter ook op een meer directe wijze betrokken worden in de kwaliteitsbewaking van evaluaties. Ze kunnen nagaan of de verzamelde evaluatiegegevens al dan niet geschikt zijn om veralgemenende uitspraken te doen over de noodzakelijke beroeps- en studievaardigheden van leerlingen. Interessant is bovendien de vraag of de externen de leerlingprestaties op eenzelfde manier beoordelen als leerkrachten. Gebruiken ze dezelfde criteria en standaarden? Zowel de validiteit als de betrouwbaarheid van een evaluatie kunnen wel varen bij een dergelijk onderzoek. Terloops willen we benadrukken dat externen niet alleen bij de kwaliteitsbewaking van evaluatie een belangrijke rol kunnen spelen. Ook bij het zoeken naar geschikte evaluatietaken en bij het verzorgen van feedback kan hun actieve bijdrage een meerwaarde hebben.

2. De rol van collega's

Ten opzichte van externen bieden collega's het voordeel dat ze makkelijk beschikbaar zijn en op een meer informele manier kunnen betrokken worden bij de kwaliteitsbewaking van evaluatie. Overleg met collega's moet garanderen dat leerkrachten onderling dezelfde criteria en standaarden hanteren en dat ze eenzelfde prestatie op eenzelfde manier beoordelen. Ook het kritisch bekijken van elkaars evaluatietaken kan op een vrij eenvoudige manier georganiseerd worden. Tevens kunnen leerkrachten gezamenlijk taken ontwikkelen en op die manier een maximale validiteit nastreven. Samenwerking speelt bovendien een rol in de efficiëntie van een evaluatie omdat dubbel werk vermeden kan worden.

3. De rol van leerlingen

De mate van objectiviteit en transparantie van een evaluatie kunnen uiteraard het best bevraagd worden bij de leerlingen. Een dergelijke kwaliteitsbewaking kan zowel voor als na de evaluatie gebeuren. Centraal staan de vragen of de leerlingen van mening zijn of de evaluatie in overeenstemming is met het curriculum en de instructie, en of leerlingen zich voldoende voorbereid weten om de evaluatie te ondergaan. Zeer informatieve informatie omtrent de validiteit van een evaluatie is bovendien te verkrijgen door leerlingen te vragen of ze van mening zijn dat ze aan de hand van een bepaalde evaluatie in staat waren (of zullen zijn) om hun eigenlijke capaciteiten aan te tonen.

Om het evalueren van een evaluatie af te ronden nog even de volgende denkoefening.

Even stilstaan... 2.6: Hoe kwaliteitsvol is de evaluatie van de evaluatie?

We spraken over het evalueren van de evaluatie. Dit impliceert dat de eisen die we vooropstellen aan de didactische evaluatie van leerlingen -mutatis mutandis- onverminderd van toepassing zijn op de evaluatie van evaluatie. Doelmatigheid en billijkheid zijn ook hier van groot belang. Enig gezond verstand en creativiteit zijn hier aan de orde. Kan u bij wijze van denkoefening aangeven tegen welke kwaliteitseisen volgende voorbeelden van evaluatie zondigen?

- ✓ Als leerkracht wil ik mijn eindexamen van het voorbije schooljaar kritisch benaderen. Daarom leg ik een kritische vragenlijst voor aan de leerlingen die nu in het volgende schooljaar zitten.
- ✓ Om de evaluatie efficiënt te laten verlopen heb ik een aantal criteria die volgens mij belangrijk zijn op een rij gezet. De leerlingen moeten op een schaal aangeven in hoeverre ze van mening zijn dat de verschillende criteria gerealiseerd werden. Leerlingen hebben wel niet de mogelijkheid om eigen criteria aan te brengen en om hun antwoord toe te lichten.
- ✓ De bevraging van de leerlingen gebeurt om praktische redenen niet anoniem.

Hoofdstuk 3: Kenmerken van alternatieve evaluatievormen

1 Inleidend

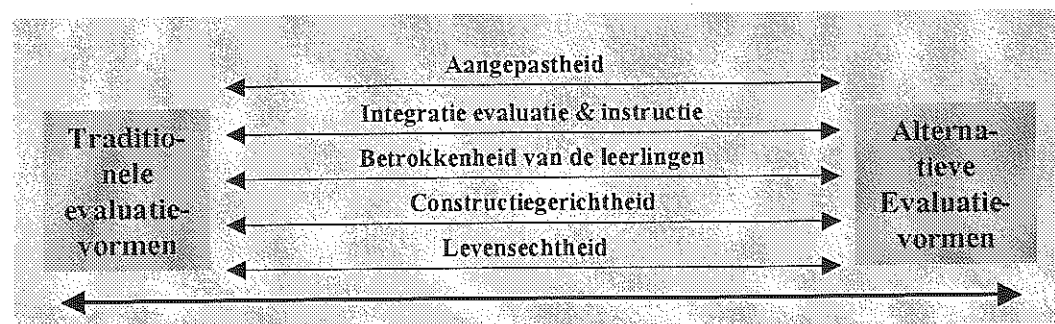
Naast de meer traditionele evaluatievormen zoals toetsen, examens en het laten uitvoeren van eenvoudige vaardigheden, gaat tegenwoordig veel aandacht naar het ontwikkelen van nieuwe evaluatievormen. Vaak groepeert men deze onder verzameltermen zoals 'alternatieve' vormen van evaluatie. Onder deze noemer vallen dan ook verschillende methoden en werkwijzen: authentieke prestatieopdrachten, simulaties, portfolio's, tijdschriften, groepsprojecten, tentoonstellingen, observaties, interviews, mondelinge presentaties, zelfevaluatie, peerevaluatie etc. In deze verscheidenheid willen we een zekere orde scheppen. Dit derde hoofdstuk geeft daartoe een algemene beschrijving van de essentiële kenmerken van dergelijke alternatieve evaluatievormen. Op die manier kan duidelijk worden welke evaluatiemethoden we als alternatief zullen bestempelen.

In vorig hoofdstuk stelden we dat een evaluatie niet alternatief hoeft te zijn om als goed te worden bestempeld. Alternatieve evaluatievormen werken immers ofwel een aantal kenmerken van goede evaluatie verder uit of voegen een aantal kenmerken toe. In de tweede paragraaf van dit hoofdstuk staan we eerst stil bij de verschillende kenmerken die maken dat we spreken van alternatieve evaluatievormen. Net door het benadrukken of toevoegen van bepaalde kenmerken zouden andere kenmerken van goede evaluatie echter in het gedrang kunnen komen. In paragraaf 3 van dit hoofdstuk zullen we daarom nagaan in welke mate alternatieve evaluatiemethoden voldoen aan de eisen van goede evaluatie. De centrale vraag is daar of de kenmerken van alternatieve evaluatievormen impliceren dat er aan andere kwaliteitseisen moet worden ingeboet.

2 Kenmerken van alternatieve evaluatievormen

De term 'alternatieve evaluatievormen' is een verzamelnaam. Verschillende evaluatiemethoden verdienen de term 'alternatieve evaluatie'. Naargelang de verschillende kenmerken die we in het volgende zullen voorstellen al dan niet vervuld worden, is er sprake van verschillende evaluatiemethoden. Niet elk kenmerk moet dus in dezelfde mate gerealiseerd worden om te kunnen spreken van een alternatieve evaluatievorm. Niettemin geldt dat een evaluatievorm naarmate ze meer tegemoet komt aan de kenmerken van alternatieve evaluatievormen beter geschikt zal zijn om de hedendaagse onderwijs- en evaluatiedoelen te realiseren. Dit impliceert dat de ideale alternatieve evaluatievorm een combinatie is van verschillende alternatieve evaluatievormen. In deel 4 bespreken we deze verschillende evaluatievormen. In het onderstaande leggen we ons toe op de kenmerken die maken dat een evaluatie als 'alternatief' kan worden bestempeld.

Alternatieve evaluatievormen kunnen niet zomaar zwart-wit worden afgezet tegenover andere evaluatiemethoden. Dit neemt niet weg dat een aantal specifieke kenmerken naar voren kunnen worden geschoven. Deze kenmerken zijn de aangepastheid aan de kenmerken van de leerlingen, de integratie van instructie en evaluatie, de betrokkenheid van leerlingen en het constructiegericht en levensechte karakter van de gebruikte evaluatietaken. Zoals uit de onderstaande figuur kan blijken, is het mogelijk elk van deze kenmerken als een continuüm voor te stellen.



Figuur 2.9: Continua richting alternatieve evaluatievormen

Zijn de verschillende kenmerken (eerder) niet van toepassing bij een specifieke evaluatievorm, dan zal deze evaluatie aanleunen bij de traditionele evaluatievormen. Indien wel wordt tegemoet gekomen aan de verschillende kenmerken spreken we van een alternatieve evaluatiemethode. Of een evaluatiemethode al dan niet alternatief is, is dus geen kwestie van alles of niets. Dit uit zich bijvoorbeeld in volgende vragen: 1. hoeveel van de kenmerken moeten vervuld zijn om van een alternatieve evaluatievorm te kunnen spreken en 2. wanneer situeert een evaluatiemethode zich voldoende rechts op het continuüm om te besluiten dat een bepaald kenmerk effectief van toepassing is?

Wellicht is het derhalve wenselijk de term alternatieve evaluatievorm eveneens als een continuüm op te vatten. We zouden dan ook kunnen stellen dat de elementen die wij in de volgende paragrafen beschrijven, kenmerkend zijn voor een ideale vorm van alternatieve evaluatie. In die ideale vorm zijn alle kenmerken van alternatieve evaluatievormen aanwezig. Hoewel evaluatievormen vaak slechts één van bovenstaande kenmerken expliciet benadrukken, zullen we ze toch als alternatief bestempelen (dit geldt bijvoorbeeld voor performance assessment en zelfevaluatie). Dit betekent dat vormen van alternatieve evaluatie (zie deel 4) niet noodzakelijk alle kenmerken van alternatieve evaluatievormen in zich dragen.

Volgende paragrafen beschrijven achtereenvolgens de vijf aangehaalde kenmerken van alternatieve evaluatievormen. Drie van deze kenmerken (betrokkenheid van leerlingen, constructiegerichtheid en levensechtheid) komen bij het beschrijven van de alternatieve evaluatievormen uitvoerig aan bod. Deze kenmerken worden hier bijgevolg enkel kort gepresenteerd. De aangepastheid aan de kenmerken van leerlingen en de integratie van evaluatie en instructie komen

later niet opnieuw aan bod. Van deze kenmerken geven we daarom hier reeds kort aan waarom waard zijn nagestreefd te worden.

2.1 De evaluatie is aangepast aan de kenmerken van de leerlingen

Ook bij alternatieve evaluatiemethoden geldt dat er niet één evaluatievorm bestaat die het meest geschikt is voor alle leerlingen. Dit kan betekenen dat het gebruik van één evaluatievorm voor alle leerlingen tot vertekeningen en invalide gegevens leidt. Er zijn immers persoonspecifieke kenmerken, naast dewelke die men wil evalueren, die van invloed kunnen zijn op de prestaties van leerlingen. Te denken valt daarbij aan cognitieve en affectieve karakteristieken en aan elementen als motivatie, tijdsinvestering, succesverwachting, doelmatigheidsbeleving of bepaalde (toevallige) zorgen en problemen tijdens de evaluatie. Het spreekt voor zich dat dergelijke aspecten een belangrijke impact kunnen hebben op de verzameling en de interpretatie van de evaluatiegegevens. Bij het uitwerken van een evaluatiemethode is het zeer belangrijk te weten dat oplossingsstrategieën van leerlingen eerder variëren ten gevolge van verschillende oplossingsstijlen dan ten gevolge van verschillen in capaciteiten. Evaluatiemethoden moeten dan ook garanderen dat leerlingen met verschillende oplossingsstijlen allemaal kunnen aantonen of ze kritisch kunnen denken en of ze al dan niet in staat zijn bepaalde problemen op te lossen (P.L.U.S., 2001).

Birenbaum (1995) beschrijft onderzoek naar effect van niet relevante karakteristieken van leerlingen. Uit deze gegevens wordt geconcludeerd dat het aangewezen is de evaluatievorm aan te passen aan de affectieve en cognitieve kenmerken van de leerling om alzo de kwaliteit van de evaluatiegegevens te verhogen. In principe zou men dus al deze kenmerken van de leerlingen moeten nagaan alvorens te evalueren. Dit is echter een te gecompliceerde en dus onmogelijke opgave in de dagelijkse klaspraktijk. Het is daarom wenselijk een alternatieve benadering te hanteren. Deze bestaat erin de leerlingen te laten kiezen welke evaluatievorm het best aansluit bij hun specifieke kenmerken. Deze benadering is gebaseerd op onderzoeksresultaten die aantonen dat leerlingen bekwaam zijn een accurate keuze te maken aangaande de moeilijkheidsgraad van een evaluatietask. Tevens blijkt dat leerlingen de evaluatievorm kunnen kiezen die resulteert in de beste weergave van hun werkelijke prestaties (Birenbaum, 1995).

De aangepastheid aan de kenmerken van de leerlingen is een verdere uitwerking van een aantal algemene kwaliteitseisen die we in het vorige hoofdstuk beschreven. Door de impact te minimaliseren van niet relevante individuele leerlingkenmerken op de verzameling en beoordeling van evaluatiegegevens, wordt immers zowel de doelmatigheid als de billijkheid van de evaluatie verhoogd. Deze eis aan alternatieve evaluatievormen wil met andere woorden benadrukken dat men individuele verschillen tussen leerlingen niet uit het oog mag verliezen en dat niet relevante verschillen het evaluatieresultaat niet mogen beïnvloeden.

Met volgende denkoefening brengen we aantal werkwijzen aan het licht die het mogelijk maken tijdens een evaluatie in te spelen op individuele kenmerken van leerlingen.

Even stilstaan... 2.7: De evaluatie aanpassen aan kenmerken van de leerling

Volgende situaties beschrijven elk een concreet evaluatieprobleem dat resulteert uit een specifiek verschil tussen leerlingen. Geef telkens aan hoe u deze problemen tussen leerlingen zou opvangen?

- ✓ Peter en Frank krijgen beiden dezelfde evaluatietaak voorgelegd. Voor Peter vormt deze taak geen probleem terwijl ze voor Frank net iets te moeilijk is. Na de evaluatie voelt Frank zich ongemakkelijk omdat de leerkracht onterecht het beeld heeft gekregen dat hij er in het geheel niets van terecht brengt.
- ✓ Karel en Rita hebben beiden op een zeer succesvolle manier materiaal verzameld over de rol van water bij het ontstaan van steden in de Middeleeuwen. Ze zijn echter verplicht om deze informatie mondeling te presenteren aan de rest van de klasgroep. Karel is daar niet goed in en zijn presentatie gaat compleet de mist in. Hij voelt zich ontevreden omdat hij de kwaliteit van zijn gegevensverzameling niet heeft kunnen demonstreren.
- ✓ Komende vrijdag hebben Ann en Petra beiden hun eindexamen voor het vak Omgangskunde. De leerlingen krijgen twee dagen vrijaf om zich voor te bereiden. Ann moet tijdens die twee dagen echter nog stage lopen omdat ze tijdens het schooljaar een paar dagen ziek is geweest. Ze vreest dat haar examenresultaat hieronder zal leiden.

Een belangrijk aandachtspunt betreft de vraag of de evaluatie(taak) is aangepast aan de voorkennis en de capaciteiten van de verschillende leerlingen. In het voorbeeld van Peter en Frank blijkt dat de taak die ze dienen op te lossen niet geschikt is om voor de beide correcte evaluatiegegevens te verzamelen. Dit omwille van verschillen in capaciteiten en/of voorkennis. Indien dergelijke verschillen bestaan moet de evaluatie hier gepast op reageren. Zo kan men leerlingen de toegang verzekeren tot een aantal informatiebronnen of men kan leerlingen - wat moeilijkheidsgraad betreft - verschillende evaluatietaken ter beschikking stellen. Op die manier kunnen leerlingen een taak kiezen waarvan de moeilijkheidsgraad hen het best in staat stelt om hun capaciteiten te demonstreren.

De evaluatiegegevens van Karel en Rita zijn om een andere reden niet aangepast. De veronderstelde manier om hun prestaties toe te lichten, houdt geen rekening met het feit dat Karel problemen heeft om zich op een vlotte manier verbaal uit te drukken. Het zou daarom beter geweest zijn om de leerlingen te laten kiezen op welke manier ze zich wensen uit te drukken (geschreven, mondeling, visueel).

Ann en Petra moeten op hetzelfde moment hun capaciteiten demonstreren. Beiden hebben echter niet dezelfde kansen gehad om zich voor te bereiden. Daarom is het in sommige gevallen wenselijk om leerlingen de kans te geven om (gedeeltelijk) zelf te beslissen wanneer ze geëvalueerd worden. Enige flexibiliteit in de timing van de evaluatie kan de kwaliteit van de evaluatiegegevens immers sterk verbeteren.

In het voorgaande benadrukten we dat evaluatie moet worden aangepast aan de kenmerken van individuele leerlingen. Dit geldt uiteraard ook voor de instructie die leerlingen ontvangen. Een optimale afstemming van het leren op de kenmerken van de leerlingen is dan ook een belangrijk doel van alternatieve evaluatievormen: zij kaderen m.a.w. binnen een leerdergecentreerd concept. De integratie van evaluatie binnen het onderwijs/leerproces speelt hierin een cruciale rol.

2.2 De integratie van evaluatie met onderwijs en leren

In de onderwijs- en evaluatieliteratuur wordt steeds meer beklemtoond dat evaluatie moet verweven worden met het onderwijs en dat evaluatie ten dienste moet worden gesteld van het leren van de leerlingen (Shepard, 2000; Struyf et al., 1999). In het eerste deel van dit handboek hebben we reeds gewezen op de evolutie van een geïsoleerde naar een geïntegreerde verhouding tussen evaluatie enerzijds en onderwijs en leren anderzijds. We stelden daar dat deze integratie kenmerkend is voor de assessmentcultuur. Deze integratie is een logisch gevolg van de veranderende functies van evaluatie. Evaluatie moet leerlingen feedback geven om het leerproces strategisch bij te sturen en leerkrachten te informeren over de noodzaak het onderwijsleerproces aan te passen. Evaluatie is dus niet alleen een instrument om het leren te documenteren, maar ook om het leren te begeleiden en te optimaliseren (McMillan, 2001).

De integratie van evaluatie en instructie lijkt op het eerste zicht een abstract en gecompliceerd iets. Toch betreft het veelal alledaagse en vaak eenvoudige klaspraktijken. Wellicht kan u volgende voorbeelden in vorm van uitspraken van leerkrachten op basis van eigen ervaringen verder aanvullen.

Figuur 2.10: Indicaties van de integratie van onderwijs en evalueren

- ✓ Ik probeer leerlingen te observeren wanneer ze effectief aan het werk zijn. Dat vertelt mij veel meer over de leerling in kwestie dan 's avonds aan mijn bureau de antwoorden op toetsen te corrigeren.
- ✓ Eigenlijk zijn de scores van leerlingen op evaluatietaken voor mij bijzaak. Als leerkracht heb ik doelen in gedachte, en ik gebruik evaluatiegegevens om uit te maken waar ik zelf aan moet werken en om na te gaan wat de leerlingen al dan niet reeds bereikt hebben. Daar probeer ik dan tijdens de instructie gepast op te reageren.
- ✓ Ik probeer bij de aanvang van een nieuwe leerinhoud zo snel mogelijk een evaluatie op poten te zetten. Op die manier kan ik leerlingen indien nodig in een aantal groepen verdelen die dan elk hun eigen begeleiding en opdrachten krijgen.
- ✓ Bij het begin van een les heb ik de gewoonte om een klassikaal evaluatiemoment in te lassen. Ervaring leert mij dat dit versterkt wat we de les voordien geleerd hebben.
- ✓ Het onderscheid tussen een leer- en een evaluatietaak is in mijn ogen onbestaande of toch minstens irrelevant. De leerlingen leren terwijl ze geëvalueerd worden en ik evalueer ze terwijl ze leren.

Deze leerkrachten zien evalueren als een onderdeel van het onderwijsleerproces. Sterker nog: het ondersteunen van het leerproces vormt de essentie van onderwijs en evaluatie is een hulpmiddel dat hiertoe moet bijdragen.

Daar alternatieve evaluatievormen gekenmerkt worden door een integratie van instructie en evaluatie gaan we dieper in op respectievelijk het waarom en het hoe van deze integratie.

2.2.1 Waarom evaluatie en instructie integreren?

Uit figuur 2.10 blijkt dat de integratie van evaluatie en instructie verschillende invullingen kan krijgen. Er worden verschillende doelen mee nagestreefd en er zijn dan ook verschillende argumenten die de integratie van instructie en evaluatie bepleiten. In essentie komen die allemaal neer op het feit dat evaluatie niet alleen moet worden opgevat als een beoordelingsinstrument, maar ook - en binnen de assessmentcultuur vooral - als een leerinstrument. Goede alternatieve evaluatie is steeds een geïntegreerd onderdeel van het onderwijsleerproces. Dit blijkt uit de nieuwe functies die aan evaluatie worden toegeschreven en aan de achterliggende doelen van deze functies.

1. Nieuwe functies van evaluatie

Evaluatie binnen een assessmentcultuur heeft een ruimere functie dan het mogelijk maken van een eindbeoordeling bij de leerlingen. Volgende drie -deels overlappende- functies (of doelen) van evaluatie maken duidelijk waarom evaluatie en instructie geïntegreerd moeten worden.

a. Evaluatie als het voortdurend opvolgen van leerlingen

Het permanent opvolgen van leerlingen is niet nieuw, ook al is het wellicht correct om te stellen dat vooral onderwijzers hier aandacht aan besteden. Doel hiervan is de leerlingen zo gericht mogelijk te begeleiden naar het bereiken van de vooropgestelde doelen. Hiertoe proberen leerkrachten allerlei zaken te bewaken die een voortdurende evaluatie en opvolging van de leerlingen vereisen. Dit uit zich in vragen als: Verstaan de leerlingen het materiaal? Aan welke leerlingen moet ik mijn vragen richten? Zijn de leerlingen aandachtig? Welke leerlingen hebben bijkomende begeleiding nodig? Bij het beantwoorden van dergelijke vragen is het belangrijk systematisch gegevens te verzamelen en niet louter uit te gaan van een intuïtief aanvoelen. Het vernieuwende van de integratie die we willen benadrukken, bestaat net uit het systematische karakter ervan.

b. Evaluatie als instrument om de instructie te optimaliseren

De integratie van evaluatie en instructie informeert leerkrachten over wat en hoe te onderwijzen. In feite betreft dit dezelfde werkwijzen en hetzelfde gebruik van evaluatiegegevens als bij het opvolgen van de leerlingen. Dit uit zich dan ook in dezelfde vragen: Verstaan de leerlingen het materiaal? Aan welke leerlingen moet ik mijn vragen richten? Zijn de leerlingen aandachtig? De insteek is evenwel anders, evaluatiegegevens worden gebruikt door de leerkracht om de instructie te optimaliseren. Een aangepaste instructie vereist een permanente gegevensverzameling en dus de integratie van evaluatie enerzijds en onderwijzen en leren anderzijds. Evaluatie kan daarnaast ook op een indirecte manier bijdragen tot een betere instructie. Wanneer leerkrachten gebinnen met het ontwikkelen van alternatieve vaardigheidstaken die zullen gebruikt worden tijdens de evaluatie komen ze vaak tot de vaststelling dat ook de instructie op een alternatieve manier zou moeten verzorgd worden. Het doorlopen van een evaluatieproces is dus tevens een instrument om het onderwijsleerproces aan een kritische blik te onderwerpen.

c. Evaluatie als bron van feedback

Ook het aangepast gebruik van feedback kan een belangrijke rol spelen in het optimaliseren van het leerproces. Meer nog: als het doel van evaluatie bestaat uit het verbeteren van leerlingprestaties -en dus niet alleen uit het documenteren ervan- dan moet de evaluatie gepaard gaan kwaliteitsvolle feedback (Wiggins, 1998). De meest effectieve feedback richt de aandacht op de vooruitgang die leerlingen al dan niet maken in het realiseren van de vooropgestelde doelen. Het is belangrijk om reeds tijdens het onderwijsleerproces, op basis van evaluatiegegevens, te duiden waar er zich nog eventuele tekorten bevinden en te identificeren waaraan de leerling nog moet werken. Het verschaffen van dergelijke feedback kan bovendien helpen een aantal competenties te ontwikkelen bij de leerlingen. Bijvoorbeeld: vaardigheden zoals het achterhalen en reflecteren over waar en waarom men fouten heeft gemaakt en het overgaan tot een planning van hoe deze fouten kunnen worden weggewerkt. Momenteel bestaat de enige feedback die leerlingen krijgen vaak uit een puntenscore als beoordeling van een toets of examen. Indien de evaluatie en instructie niet geïntegreerd verlopen, kan de leerling met dergelijke informatie weinig aanvangen. Het geëvalueerde leerstofonderdeel is immers afgelopen en de leerling heeft de kans niet meer om zich verder te ontwikkelen. Bij een geïntegreerde evaluatie en instructie ligt dat anders: de leerlingen kunnen met de feedback die ze krijgen verder aan de slag.

Alternatieve evaluatiemethoden hebben niet alleen het doel na te gaan welke onderwijsdoelen de leerlingen al dan niet bereikt hebben. In de vorige paragrafen stelden we dat men op basis van de evaluatiegegevens ook de manier waarop men deze doelstellingen tracht te bereiken, kan afstemmen op individuele leerlingkarakteristieken. Op die manier kan evaluatie het leerproces van leerlingen optimaliseren én kan het de planning en uitvoering van de instructie ondersteunen. Wil je inspelen op specifieke kenmerken van het leerproces van leerlingen en wil je het leerproces bijsturen, dan is een koppeling van onderwijzen en evalueren vereist. Het voortdurend opvolgen van leerlingen, het gebruik van evaluatie als instrument om de instructie te optimaliseren en het verzorgen van goede feedback hebben echter elk hetzelfde achterliggende doel voor ogen, namelijk het optimaliseren van het leerproces. Dit betekent dat het meest fundamentele argument om evaluatie en instructie te integreren, neerkomt op het gebruik van evaluatie als een leerinstrument.

2. Evaluatie als een leerinstrument

Met het gebruik van evaluatie als leerinstrument bedoelen we twee verschillende zaken die beide onlosmakelijk verbonden zijn met de integratie van evaluatie en instructie. Enerzijds kan evaluatie bijdragen tot een optimalisatie van het individuele leerproces en anderzijds kunnen leerlingen zaken leren tijdens het evaluatieproces op zich. Volgende paragrafen verduidelijken deze twee aspecten.

a. Optimalisering van het individuele leerproces

In de vorige paragraaf beschreven we een aantal nieuwe functies van evaluatie. Het doel van elk van deze functies is het optimaliseren van het leerproces van de geëvalueerde leerling.

Op welke manier deze evaluatiefuncties het leerproces kunnen optimaliseren willen we aantonen aan de hand van twee voordelen die een dergelijk gebruik van evaluatiegegevens met zich meebrengt, namelijk het uitlokken van taakgerichtheid en het nastreven van een hoger verwerkingsniveau.

- *Taakgerichtheid uitlokken*

De doelen die leerlingen zichzelf stellen blijken van groot belang voor het verloop van het leerproces. In de onderwijsliteratuur wordt in dat opzicht vaak een onderscheid gemaakt tussen taakgerichte en prestatiegerichte leerlingen. Taakgerichte leerlingen zijn gericht op het verwerven van kennis en vaardigheden, en op het nastreven van leerdelen. Deze leerlingen willen zichzelf bekwamen. Prestatiegerichte leerlingen streven ernaar beter te zijn dan de andere leerlingen en willen dus prestatiedoelen realiseren. Het onderscheid tussen beiden is belangrijk daar gebleken is dat het nastreven van leerdelen de meest gunstige effecten heeft op motivatie, cognitie en affect (Struyf et al., 1999). De verschillen in doelgerichtheid zijn geen vaststaand persoonlijkheidskenmerk maar variëren naargelang de onderwijscontext. Afhankelijk van de onderwijsleeromgeving zullen leerlingen dus meer of minder door één van beide doelgerichtheden gekenmerkt worden. Ook de evaluatievorm die men hanteert, speelt hierbij een rol. Struyf et al. (1999, p. 13) beschrijven het belang van de evaluatie als volgt: "Een leerkracht die enkel oog heeft voor het leerresultaat of de leerprestaties van leerlingen -ongeacht de beginsituatie van de leerling, zijn inzet of het leerproces dat hij reeds doormaakte- en die enkel leerlingen die in verhouding tot de klasgroep hoge punten behalen, prijst en loont, zal bij de leerlingen een prestatiegerichtheid uitlokken. Het doel van leren is dan presteren en mislukkingen vermijden. Succes betekent beter zijn dan de anderen. Een leerkracht die daarentegen het leerproces evalueert en die fouten niet beschouwt als een tekort aan intelligentie, maar als een onderdeel van het leerproces, zal bij de leerlingen taakgerichtheid uitlokken. Het doel van leren is dan beheersing. Succes is gelijk aan vooruitgang ten opzichte van een bepaald criterium of van vroegere prestaties. Leerlingen die een inspanning hebben geleverd worden hiervoor geprezen en beloond. Om hoge mate van taakgerichtheid te realiseren bij leerlingen moeten het leer- en het evaluatieproces dus op elkaar afgestemd en geïntegreerd worden. Men mag beide elementen dus niet van elkaar loskoppelen.

- *Een hoger verwerkingniveau nastreven*

Bij het beschrijven van de assessmentcultuur in het eerste deel van dit boek benadrukten we dat de huidige onderwijsbenaderingen een diepgaand cognitief verwerkingsniveau trachten te realiseren. We stelden dat het verwerkingsniveau samenhangt met bepaalde persoonsvariabelen en kenmerken van de leeromgeving. Eén aspect hiervan is de evaluatiepraktijk van leerkrachten. Een evaluatiepraktijk die geïntegreerd is in het onderwijsleerproces en die is afgestemd op individuele leerlingkenmerken kan een aantal garanties bieden voor het effectief realiseren van diepteverwerking (Struyf et al., 1999). Zo zal het stellen van evaluatieve vragen van cognitief hogere orde (zoals verbandsvragen, inzichtsvragen, e.d.) tijdens het onderwijsleerproces de leerlingen stimuleren om

bij het instuderen van leerstof eveneens gebruik te maken van verwerkingsstrategieën van een hogere orde. De manier waarop leerlingen de leerstof verwerken, is immers afhankelijk van de verwachting die ze hebben over wat hen tijdens de evaluatie gevraagd zal worden. Bijvoorbeeld: aan welke eisen of verwachtingen moet ik voldoen. Die verwachtingen zijn niet voor alle leerlingen gelijk en moeten dus worden afgestemd op individuele leerlingkenmerken. Dat kan alleen als evaluatie en instructie geïntegreerd verlopen.

- b. *Leren tijdens de evaluatie*

Traditioneel is leren losgekoppeld van evalueren. Tijdens de evaluatie gaat men na wat leerlingen bereikt hebben en daarmee is de kous af. Deze benadering gaat echter voorbij aan de leerkansen die een evaluatie de leerlingen kan bieden. Dit blijkt bijvoorbeeld uit de volgende doelstelling die we voor leerlingen BSO - Verzorging vooropstelden: De leerlingen kunnen positieve en negatieve eigenschappen van de eigen persoon als verzorgende opnoemen. Net door zichzelf te evalueren kunnen leerlingen zich aldus belangrijke vaardigheden eigen maken. Goed ontwikkelde evaluatietaken documenteren dus niet alleen de vooruitgang die leerlingen boeken; de evaluatie kan leerlingen ook effectief helpen deze doelen te bereiken. Evaluatietaken zijn in dat geval tevens leertaken. Niet alleen het betrekken van leerlingen in de eigenlijke beoordeling leidt tot leerkansen, ook andere aspecten van het evaluatieproces kunnen maken dat leerlingen belangrijke zaken opsteken. Hierbij denken we bijvoorbeeld aan het ontwikkelen en selecteren van evaluatiecriteria en aan het op zoek gaan naar evaluatietaken die geschikt kunnen zijn om de vooropgestelde doelen te evalueren.

Het gebruik van evaluatie als een leerinstrument is in principe de voornaamste reden om evaluatie en instructie te integreren. Andere motieven komen hier onrechtstreeks immers eveneens op neer. Dat betekent dan ook dat de belangrijkste argumenten in de vorige paragrafen zijn aangehaald. Om ons pleidooi inzake de integratie van evaluatie en instructie af te ronden, richten we ons even niet op de voordelen van de integratie, maar wel op de nadelen van een eventuele isolatie.

3. *De negatieve impact van geïsoleerde instructie en evaluatie*

De integratie van instructie en evaluatie wordt tevens bepleit op basis van de negatieve, ongewenste implicaties van een niet-geïntegreerde instructie en evaluatie. Deze invalshoek is interessant om de eigen evaluatiepraktijk even kritisch te bekijken. Dit beogen we met volgende denkoefening.

Even stilstaan... 2.8: Kritieken op geïsoleerde instructie en evaluatie

Neem even een concrete onderwijssituatie in gedachte die u zelf hebt ingericht of ondergaan? Stel uzelf hierbij de vraag in welke mate de volgende kritieken op en negatieve gevolgen van een geïsoleerde evaluatie en instructie op deze situatie van toepassing zijn.

- ✓ Leerlingen krijgen enkel te horen wat de kwaliteit is van hun huidige prestaties zonder een goed zicht te verwerven op hun eigen leerproces. Leerlingen hebben geen beeld van welke vooruitgang ze nog zouden kunnen boeken.
- ✓ Hetzelfde geldt voor de leerkracht. De evaluatie wordt enkel gebruikt als afsluiter van een les-eenheid zodat het voor de leerkracht onmogelijk is om zijn instructie aan te passen op basis van de verkregen evaluatiegegevens. Na de evaluatie is de leerkracht immers alweer met andere leerstof bezig.
- ✓ De informatie die leerlingen krijgen heeft voornamelijk betrekking op de vraag of ze beter of slechter presteren dan hun klasgenoten. Of leerlingen in vergelijking met hun eigen vroegere prestaties al dan niet beter presteren komen ze niet te weten.
- ✓ Leerlingen weten niet goed wat er van hen verwacht wordt. Pas wanneer ze de eindevaluatie achter de rug hebben, wordt hun duidelijk wat de doelen waren die de leerkracht wilde bereiken.

Met deze punten van kritiek willen we allerm minst beweren dat leerkrachten momenteel in het geheel slecht scoren wat de integratie van evaluatie en instructie betreft. Verschillende leerkrachten leveren immers reeds mooie voorbeelden van deze integratie. Deze presenteren we in de volgende paragraaf.

2.2.2 Hoe evaluatie en instructie integreren?

Het verduidelijken van het waarom van de integratie brengt ons bij de vraag hoe evaluatie en instructie kunnen geïntegreerd worden. Ook al klinken al deze termen goed in de oren, daarmee is evenwel niet meteen duidelijk hoe een dergelijke integratie er in de dagelijkse klaspraktijk kan uitzien. Toch is het in wezen eenvoudig. Het integreren van evaluatie en instructie betekent zoveel als het voortdurend verzamelen van informatie over de leerlingen (Rudman, 1989). Tijdens de instructie vervult een leerkracht als het ware twee taken: hij verzorgt de instructie en evalueert voortdurend de vooruitgang van het leer- en instructieproces. Het integreren van evaluatie en instructie impliceert niet zozeer dat je als leerkracht volkomen nieuwe evaluatie-instrumenten moet gaan ontwikkelen, maar wel dat de evaluatiegegevens anders moeten gebruikt worden. Dit kan op de zoals hoger beschreven manier en met de nieuwe functies in het achterhoofd. Om evaluatie en instructie te integreren is het dus niet noodzakelijk om zogenaamde alternatieve evaluatievormen te hanteren. De resultaten van een klassieke meerkeuzetoets kunnen evenzeer leiden tot een geslaagde integratie, zolang de evaluatiegegevens maar voor de vooropgestelde doelen en functies gebruikt worden. Dit neemt uiteraard niet weg dat wij ervoor pleiten om zoveel mogelijk tegemoet te komen aan de kenmerken van alternatieve evaluatievormen. Om dit op een succesvolle manier te doen, kunnen een aantal voorwaarden beschreven worden. Die zijn bijvoorbeeld het feit dat er duidelijkheid moet bestaan over de aard van de beslissingen die kunnen volgen op de evaluatie en het feit dat de evaluatie zodanig moet georganiseerd worden dat de leerlingen er door gemotiveerd worden om te blijven leren. Dergelijke elementen worden elders uitgewerkt.

2.3 Betrokkenheid van leerkrachten en studenten

Niet alleen de rol van studenten is aan verandering toe wanneer men alternatieve evaluatiemethoden wil hanteren; ook de instelling van leerkrachten moet gekenmerkt worden door een mentaliteitsverandering. Toch is net het feit dat leerlingen actiever worden betrokken bij de evaluatie kenmerkend voor alternatieve evaluatievormen. Wanneer men leerlingen wil opvoeden tot autonome lerenden moeten ze ook leren zichzelf te evalueren. Op die manier wordt evaluatie de gedeelde verantwoordelijkheid van de leerkracht én de leerlingen. De rol van de leerlingen kan bij alternatieve evaluatievormen zowel voor, tijdens als na de eigenlijke evaluatie actiever worden ingevuld.

Om de betrokkenheid van leerlingen als kenmerk van alternatieve evaluatievormen te duiden, geven we in de volgende alinea's kort aan welke rol leerlingen kunnen spelen in de verschillende fases van het evaluatieproces. Merk hierbij op dat leerlingenbetrokkenheid –net als de andere kenmerken van alternatieve evaluatievormen– geen kwestie is van alles of niets. Bij alternatieve evaluatievormen gaat het erom de participatie van leerlingen in het evaluatieproces vorm te geven naargelang van de voorgestelde onderwijs- en evaluatiedoelen. Dit komt evenwel vaak neer op een zo groot mogelijke betrokkenheid van de leerlingen.

Alvorens exemplarisch de potentiële rollen van leerlingen te illustreren staan we stil bij de eigen evaluatiepraktijk.

Even stilstaan... 2.9: De betrokkenheid van leerlingen in de eigen evaluatiepraktijk

Neem even een concrete evaluatiesituatie in gedachte die u zelf hebt ingericht of ondergaan? Beantwoord met deze situatie in het achterhoofd volgende vragen over de betrokkenheid van leerlingen. Op die manier achterhalen we in welke mate voor leerlingen een actieve rol is weggelegd in de evaluatie.

- ✓ Hebben leerlingen de kans om mee te beslissen wat er van hen op het einde van een lessenreeks van hen verwacht wordt?
- ✓ Hebben leerlingen de mogelijkheid om te kiezen hoe ze zullen demonstreren of ze die verwachtingen al dan niet kunnen waarmaken?
- ✓ Gebeurt het dat leerlingen hun eigen prestaties moeten beoordelen? En die van medeleerlingen?
- ✓ Gebeurt het dat leerlingen feedback geven op de prestaties van anderen?
- ✓ Hoe belangrijk is reflectie op het eigen werk –volgens u– om tot betere leerlingenprestaties te komen?

Uit deze denkoefening blijkt dat leerlingen op verschillende manieren kunnen betrokken worden in het evaluatieproces en dit zowel voor, tijdens als na de eigenlijke evaluatie.

1. Betrokkenheid van leerlingen voor de evaluatie.

Anders dan bij de traditionele evaluatiebenaderingen zijn bij alternatieve evaluatiemethoden de beoordelingscriteria waarop de beoordeling gebaseerd zal worden meestal bekend bij de leerlingen. Leerlingen hebben een duidelijk zicht op wat de te bereiken doelstellingen zijn en op de criteria waaraan een prestatie moet voldoen om een bepaalde beoordeling te krijgen. Leerlingen kunnen bovendien betrokken worden in het ontwikkelingsproces van evaluatiecriteria en -instrumenten. Wanneer leerlingen weten welke de criteria zijn en wat ze eigenlijk moeten bereiken, gaan ze ook gemotiveerder en beter presteren (Dochy & Segers, 1999; Lazzaro, 1996). Leerlingen die de evaluatiecriteria kennen en ze kunnen gebruiken voor de eigenlijke evaluatie zullen hun rol immers anders invullen. Niet alleen doelen en criteria kunnen echter de betrokkenheid van leerlingen voor de evaluatie aanmoedigen. Men kan leerlingen bijvoorbeeld ook betrekken in beslissingen over de wijze waarop en over wanneer geëvalueerd zal worden.

2. Betrokkenheid van leerlingen tijdens de evaluatie.

Ook tijdens de evaluatie is de rol van leerlingen bij alternatieve evaluatievormen duidelijk verschillend van die tijdens de traditionele evaluatievormen. In de eerste plaats leren leerlingen over hun eigen werk en leerproces na te denken. Doel is leerlingen te leren zichzelf systematisch te evalueren; komen tot *zelfevaluatie* dus. Leerlingen ontwikkelen eigen criteria, beoordelen zichzelf en leren hun eigen sterke en zwakke punten op te sporen om in functie hiervan hun verdere leerproces te plannen. Wanneer de leerlingen actief betrokken worden bij de eigenlijke evaluatie, maar door de leerkracht gestuurd en gereguleerd worden, spreken we van *collaboratieve evaluatie* (of co-evaluatie). In dat geval bepaalt de leerkracht samen met de leerlingen welke de onderwijsdoelen zijn, aan welke criteria de leerlingenprestaties moeten tegemoet komen en hoe de onderwijsdoelen geëvalueerd zullen worden. Een andere vorm van evaluatie die de activiteit van leerlingen tijdens de evaluatie nastreeft, is *peerevaluatie*. In dat geval discussiëren leerlingen onderling over beoordelingscriteria en evalueren ze elkaars werk (Topping, 1998). Leerlingen krijgen de kans om hun klasgenoten te evalueren. Die samenwerking en de ervaringen die leerlingen ermee opdoen, helpen leerlingen verantwoordelijkheid te nemen voor hun eigen leren (Kerka, 1995; Janssens et al., 2000).

3. Betrokkenheid van leerlingen na de evaluatie.

Na de evaluatie is er geen sprake van eenrichtingsverkeer van feedback van de leerkracht naar de leerling. De leerkracht en de leerling (en eventueel klasgenoten) reflecteren samen over de evaluatieresultaten en zoeken manieren om verdere vooruitgang te boeken. Feedback is dus niet alleen afkomstig van de leerkracht maar ook van de leerling zelf en van de medeleerlingen. Ook met het oog op de meta-evaluatie kan het waardevol zijn de leerlingen te activeren na de eigenlijke evaluatie. Diegenen die geëvalueerd worden, kunnen immers belangrijke informatie verstrekken over hoe zij de doelmatigheid en billijkheid van de evaluatie percipiëren. Zo kunnen zij zelf een belangrijke bijdrage leveren in het verhogen van de kwaliteit van de evaluatie.

Ook al beschreven we zoniet een aantal werkwijzen die de inbreng van de lerende bij de evaluatie kunnen vergroten, toch willen we benadrukken dat de onderwijsgevende en de leerlin-

gen/studenten moeten samenwerken aan de evaluatie. Op die manier wordt de evaluatie zowel de verantwoordelijkheid van de leerkracht als die van de leerlingen. Dit betekent dat de evaluatiepraktijk niet meer volledig in handen ligt van de leerkracht en dat de leerlingen zelf mede verantwoordelijk zijn voor de evaluatie van hun eigen leerproces. Daarmee is de kern van de betrokkenheid van leerlingen als kenmerk van alternatieve evaluatievormen toegelicht. Bij het bespreken van zelf-, peer- en collaboratieve evaluatie in deel 4 gaan we dieper in op de invulling van de leerlingenbetrokkenheid.

2.4 Kennisconstructie is het uitgangspunt, niet kennisreproductie

Moderne leertheorieën stellen dat de onderwijsinhoud, het onderwijsleerproces en de evaluatie onlosmakelijk met elkaar verbonden zijn. Alternatieve evaluatievormen meten dan ook niet zozeer de reproductie van kennis, maar richten zich eerder op het evalueren van vaardigheden, metacognitie en attitudes. In algemene termen kunnen we stellen dat evaluatietaken overeen moeten komen met de doelstellingen die men voor het onderwijsleerproces vooropstelde. Het komt er dan ook op aan leerlingen kansen te bieden om hun kennis en vaardigheden te demonstreren en de evaluatie dus niet te beperken tot het herinneren van ingestudeerde feiten. Een dergelijke benadering weerspiegelt zich uiteraard in de gebruikte evaluatieopdrachten en -taken. Zo verwacht men bijvoorbeeld van leerlingen om een essay te schrijven, om in groep een wetenschappelijk experiment uit te voeren, om het antwoord op een probleem schriftelijk te beargumenteren of om een portfolio bij te houden van hun prestaties (O'Neil, 1992). Al deze evaluatievormen zijn verschillende manieren om de performantie van leerlingen te evalueren. Kenmerkend voor vaardigheidstaken is dat ze van de leerlingen verwachten dat ze een antwoord of een product creëren waarmee ze hun kennis en vaardigheden kunnen aantonen. Doorheen de evaluaties is er verder meestal een nadruk op het proces dat aan de grondslag ligt van het oplossingsproces terwijl de correctheid van de oplossing van secundair belang is (P.L.U.S., 2001).

Wat en hoe zou moeten beoordeeld worden, plaats men in de evaluatieliteratuur onder de noemer vaardigheidsevaluatie (of in het Engels performance assessment). De term vaardigheidsevaluatie zegt dus niet alleen iets over de inhoud maar ook over de vorm van de evaluatie. Vaardigheidsevaluatie wordt gedefinieerd als een systematische poging om de toepassingsvaardigheden van de lerende te meten. De vraag is dus of de lerende in staat is om reeds verworven kennis te gebruiken bij het oplossen van nieuwe problemen. Door het geven van vaardigheidstaken wil men leerlingen/studenten laten demonstreren dat ze bepaalde taken kunnen uitvoeren en dat ze de vooropgestelde vaardigheden beheersen. Met andere woorden: kennisconstructie is het uitgangspunt, niet kennisreproductie.

Bij het bespreken van vaardigheidsevaluatie als een vorm van alternatieve evaluatie (zie deel 4) gaan we uitgebreid in op constructiegerichtheid als kenmerk van vaardigheidstaken en op de manier waarop dergelijke taken kunnen ontwikkeld worden.

2.5 Gebruik van authentieke of levensechte prestatiesituaties

In moderne curriculumtheorieën wordt het belang van het betrekken van leerlingen in authentieke, realistische taken benadrukt. Deze taken zijn meer motiverend en garanderen een betere transfer van het geleerde dan de meer traditionele, gedecontextualiseerde academische taken (Herman et al., 1992; Arter & Bond, 1996; Winograd & Perkins, 1996). De aandacht voor de authenticiteit (echtheid) is er dan ook gekomen omdat in het verleden bleek dat de evaluatievragen, -taken en -opdrachten vaak veraf stonden van de werkelijkheid en daardoor als weinig zinvol werden ervaren (Janssens et al., 2000). Authentieke evaluatietaken hebben daarom tot doel delen van de realiteit en probleemoplossingsituaties te voorzien waarin leerlingen toepassingen van relevante (componenten van) vaardigheden en kennis kunnen demonstreren. Het gebruik van authentieke taken vereist dat leerlingen actief construeren en denken vanuit werkelijkheidsgetrouwe situaties en simulaties. Om authentiek te zijn, moet een taak dus in de eerste plaats constructiegericht zijn terwijl een bijkomend kenmerk nagestreefd wordt: de taken en opdrachten moeten aansluiten bij de activiteiten die de leerling in het latere (beroeps)leven zal moeten uitvoeren. Hierbij wordt de leerlingen gevraagd kennis, vaardigheden, competenties te demonstreren op een manier en in een context die door de leerlingen zelf als zinvol wordt ervaren (Dochy & Segers, 1999; ASCD, 1996). Om van levensechte taken te kunnen spreken moet men naast de inhoud van de evaluatie eveneens aandacht besteden aan de organisatie, aan het 'hoe' van de evaluatie. Ook de wijze waarop leerlingen geëvalueerd worden, moet aansluiten bij het dagelijkse (professionele) leven.

Kenmerkend voor authentieke taken is tevens dat ze vaak niet mooi samenvallen met de vak-scheiding in het secundair onderwijs. Levensechte situaties vereisen immers dat informatie wordt verzameld uit verschillende disciplines en perspectieven. Herman et al. (1992) zien hierin een argument om het onderwijsleerproces vakoverschrijdend te organiseren of om verschillende vakken samen te evalueren. Vakoverschrijdende evaluatietaken hebben bovendien het voordeel minder tijds- en kostintensief te zijn en alzo een kans te bieden om in het algemeen een betere inhoudsvaliditeit te garanderen.

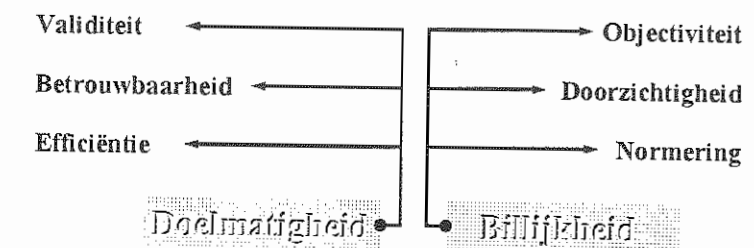
Bij het bespreken van authentieke evaluatie als een alternatieve evaluatievorm en bij het beschrijven van het evaluatieproces besteden we uitvoeriger aandacht aan de argumenten die het gebruik van authentieke evaluatietaken bepleiten en aan de manier waarop dergelijke taken kunnen worden ontworpen.

2.6 Concluderend: kenmerken van alternatieve evaluatievormen

In de beschrijving van de kenmerken van alternatieve evaluatievormen kunnen we twee delen onderscheiden. We stelden dat ook alternatieve evaluatievormen zoveel mogelijk moeten voldoen aan de eisen die worden vooropgesteld aan evaluatie in het algemeen. Daarnaast beschreven we een aantal kenmerken die typerend zijn voor de zogenaamde alternatieve evaluatievor-

men. De onderstaande figuur brengt schematisch deze beide aspecten van goede alternatieve evaluatie samen.

Figuur 2.11: Kenmerken van goede alternatieve evaluatie



Kenmerken goede alternatieve evaluatie

- ✓ Kennisconstructie is het uitgangspunt, niet kennisreproductie
- ✓ De evaluatie is aangepast aan de kenmerken van de leerlingen
- ✓ De integratie van evaluatie met onderwijzen en leren
- ✓ Betrokkenheid van leerkrachten en studenten
- ✓ Gebruik van authentieke of levensechte situaties
- ✓ Zowel het proces als het product van het leren zijn belangrijk

Bij deze voorstelling dringt zich een opmerking op, die zowel geldt voor de levensechtheid van evaluatietaken als voor andere kenmerken van alternatieve evaluatievormen. De authenticiteit van een evaluatietask moet niet uitsluitend gezien worden als een kenmerk van alternatieve evaluatievormen. In hoofdstuk 2 van dit deel stelden we reeds dat authenticiteit eveneens naar voren wordt geschoven als een kwaliteitseis om te kunnen spreken van een doelmatige, c.q. valide evaluatie in het algemeen. In een kanteling naar een assessmentcultuur is dit een kwestie van invalshoek: de inhoud en het belang van het kenmerk zijn gelijk ongeacht ze worden vooropgesteld als kenmerk van een goede evaluatie of extra benadrukt als een kenmerken van alternatieve evaluatievormen. Hetzelfde kan gezegd worden over de constructiegerichtheid als kenmerk van een alternatieve evaluatietask omdat constructiegerichtheid in de meeste gevallen de validiteit van een evaluatiebeoordeling eveneens ten goede komt. Als een leerkracht tracht een evaluatie te ontwerpen die voldoet aan doelmatigheids- en billijkheidseisen komt hij vaak onbewust uit bij een evaluatiepraktijk die (een aantal) kenmerken van alternatieve evaluatie vertoont. Het onderscheid tussen kenmerken van goede evaluatie in het algemeen en van alternatieve evaluatievormen in het bijzonder is dus niet altijd strikt te maken. Alternatieve evaluatievormen zijn met andere woorden vaak succesvolle uitwerkingen van de kenmerken van goede evaluatie. Toch kan ook het tegenovergestelde waar zijn. Het zou immers kunnen dat door bepaalde kenmerken van alternatieve evaluatievormen te realiseren andere kenmerken van goede evaluatie in het gedrang komen. De vraag of dit zo is, staat centraal de volgende paragraaf.

3 Voldoen alternatieve evaluatievormen aan de eisen van goede evaluatie?

Ter afsluiting van dit hoofdstuk gaan we na in welke mate alternatieve evaluatiemethoden voldoen aan de eisen van goede evaluatie. De centrale vraag is of bepaalde kenmerken van alternatieve evaluatievormen impliceren dat de algemene kwaliteitseisen logischerwijze ook vervuld zijn dan wel dat er aan andere kwaliteitseisen moet worden ingebonden.

Bij het doornemen van de kenmerken van goede evaluatie enerzijds en de kenmerken van alternatieve evaluatievormen anderzijds bleek alvast dat tussen beide een zekere overlap bestaat. Getuige hiervan tevens de uitkomst van volgende denkoefening.

Even stilstaan... 2.10: Overlappende evaluatiekenmerken

Volgende vragen en antwoorden illustreren de overlap tussen algemene kwaliteitseisen en de kenmerken van alternatieve evaluatievormen. Probeer steeds even voor uzelf een antwoord te formuleren alvorens verder te lezen.

- ✓ Op welke manier kan men als evaluator ervoor zorgen dat men precies beoordeelt wat men bedoeld heeft te evalueren (cf. het validiteitsvraagstuk)? *Zou u ook niet adviseren leerlingen zoveel mogelijk onder te brengen in authentieke contexten waarin ze constructiegerichtere taken dienen uit te voeren? Op die manier kunnen de beoogde gedragingen en vaardigheden zo valide mogelijk geobserveerd worden.*
- ✓ Op welke manier kan men er als evaluator voor zorgen dat een evaluatiemethode voldoende doorzichtig is? En verder: dat de leerling alle informatie en kansen krijgt die hij nodig heeft om te zorgen voor een optimale voorbereiding op en een adequate uitvoering van de evaluatieopdracht (cf. objectiviteits- en transparantievraagstuk)? *Zou u ook niet adviseren om de evaluatie zoveel mogelijk aan te passen aan individuele kenmerken van de verschillende leerlingen en om leerlingen actief bij het ontstaan en uitvoeren van de evaluatie te betrekken?*
- ✓ Op welke manier kan men als evaluator ervoor zorgen dat een evaluatie vrij is van meetfouten, dat de beoordeling consistent, nauwkeurig en reproduceerbaar is (cf. het doelmatigheidsvraagstuk). *Zou u ook niet adviseren om zoveel mogelijk vooraf te duiden welke leeruitkomsten belangrijk zijn, om duidelijk de criteria voor een goede prestatie te omschrijven en om standaarden te ontwikkelen? Hetzelfde advies kan bovendien een belangrijke rol spelen bij het garanderen van de objectiviteit van de evaluatie en bij het garanderen van een goede normering.*

Het lijkt erop dat wanneer een aantal criteria voor goede evaluatie zo goed mogelijk nagestreefd worden, er sowieso sprake is van specifieke alternatieve evaluatievormen. Dit is ook enigszins zo. Leerkrachten wier evaluatie tegemoet komt aan de vooropgestelde eisen zullen vaak niet ver afstaan van het gebruik van wat wij alternatieve evaluatievormen noemen. Toch is deze conclusie om een aantal redenen niet volledig gerechtvaardigd. Ten eerste omdat er goede toetsen kunnen ontworpen worden die wel voldoen aan de vooropgestelde algemene kwaliteitseisen zonder kenmerken te dragen van alternatieve evaluatievormen. Ten tweede omdat om van een

ideale alternatieve evaluatiemethode te kunnen spreken nog bijkomende eisen moeten in acht worden genomen (zoals de integratie van instructie en evaluatie en de aandacht voor het evaluatieproces). De derde reden brengt ons terug bij het uitgangspunt van zoëven, namelijk de vraag of alternatieve evaluatievormen sowieso tegemoetkomen aan de eisen van goede evaluatie. Bepaalde kenmerken van alternatieve evaluatievormen zouden integendeel immers een aantal algemene kwaliteitseisen kunnen ondermijnen.

Hieronder staan we stil bij de kenmerken van alternatieve evaluatievormen die mogelijks de algemene kwaliteitseisen kunnen bedreigen. We gaan tevens na hoe in dat geval alternatieve evaluatiemethoden alsnog zoveel mogelijk aan de hoger beschreven eisen kunnen voldoen.

3.1 De doelmatigheid en billijkheid van alternatieve evaluatievormen

Alternatieve evaluatiemethoden lijken op het eerste zicht zeer valide (begripsvaliditeit en inhoudsvaliditeit) omdat ze de eigenlijke gedragingen en vaardigheden van leerlingen observeren en beoordelen (in tegenstelling tot tests waar men zich baseert op indicatoren van bepaalde vaardigheden). We mogen alternatieve evaluatievormen evenwel niet beoordelen op basis van wat op het eerste zicht lijkt. Het is op basis van empirische evidentie dat de validiteit van alternatieve evaluatiemethoden in de toekomst zal moeten worden aangetoond (Birenbaum, 1995). Blijkens volgende problemen met de validiteit van alternatieve evaluatievormen wordt alvast wel duidelijk dat bij de veelbelovendheid van deze evaluatievormen toch ook enkele kanttekeningen moeten worden gemaakt.

- *Aangaande inhoud- en domeinspecificiteit.* De taken bij alternatieve evaluatiemethoden zijn vaak weinig afgelijnd en slechts in mindere mate gestructureerd. De vraag is dan ook over welke inhoudsdomeneinen men uitspraken kan doen op basis van de verkregen evaluatiegegevens. Het is bijgevolg belangrijk bij alternatieve evaluatievormen de inhoudelijke domeinen die men wil evalueren duidelijk te specificeren en de evaluatie-instrumenten binnen de evaluatie zodanig te ontwerpen dat alle domeinen opgenomen worden (Birenbaum, 1995).
- *Aangaande 'test pollution'.* Wanneer men bepaalde inferenties (al dan niet met verstrekende gevolgen) wil formuleren op basis van de resultaten van een evaluatie dan is validiteit een zeer belangrijk gegeven. Een concreet probleem dat zich bij alternatieve evaluatievormen kan voordoen, is de zogenaamde 'test pollution' (Birenbaum, 1995). Enerzijds is er empirische evidentie dat alternatieve evaluatiemethoden taken hanteren die hogere cognitieve vaardigheden kunnen meten. Anderzijds blijkt dat leerlingen deze taken soms memoriseren zonder de bedoelde vaardigheden meester te worden. Op die manier verbeteren de testresultaten terwijl de prestaties eigenlijk achteruitgaan. Om te verzekeren dat alternatieve evaluatievormen effectief tot de bedoelde gevolgen leiden, suggereert Birenbaum (1995) dat men zich ervan moet vergewissen dat de resultaten niet het gevolg zijn van het memoriseren maar dat ze indicatief zijn voor het cognitieve

proces dat noodzakelijk geacht werd om de evaluatietask succesvol te volbrengen. Het kan hierbij zinvol zijn tijdens de evaluatie aandacht te hebben voor de tijdsduur die het oplossen van de taak bij de leerlingen vergt, voor de hulp die hen geboden wordt, voor gegevens over het oplossingsproces zelf en voor zelfreflectie door de leerlingen.

Ook wat de betrouwbaarheid van alternatieve evaluatievormen betreft, zijn een aantal opmerkingen vermeldenswaardig. In vergelijking met de traditionele tests lijken alternatieve evaluatievormen minder te voldoen aan de eisen van interbeoordelaarsbetrouwbaarheid en herhaalbaarheid. Bij een portfolio is dit bijvoorbeeld mede toe te schrijven aan de weinig gestandaardiseerde taken en de aanzienlijke vrijheid voor de leerlingen omtrent wat al dan niet in de portfolio op te nemen. De interbeoordelaarsbetrouwbaarheid en herhaalbaarheid zouden met andere woorden tekort schieten. Sommige auteurs gaan zover te stellen dat alternatieve evaluatiemethoden in het geheel niet kunnen tegemoet komen aan de eisen van betrouwbaarheid (Birenbaum, 1995). Uit dergelijke stellingen onthouden wij dat de betrouwbaarheid van alternatieve evaluatievormen nauwlettend in de gaten moet worden gehouden. Mits het evaluatieproces dat we in deel 3 toelichten zorgvuldig doorlopen wordt, kan de betrouwbaarheid van alternatieve evaluatievormen ons inziens immers best wel gewaarborgd worden. In concreto denken we hierbij aan het duidelijk expliciteren van verwachte leeruitkomsten, evaluatiecriteria, rubrieken en evaluatiestandaarden en aan het zorgvuldig ontwikkelen van evaluatie-instrumenten. Illustratief rapporteren we hier een aantal bijkomende werkwijzen die de betrouwbaarheidsproblemen kunnen voorkomen. Zo zou men het oordeel kunnen laten afhangen van een onderlinge bespreking van individuele evaluaties. Verschillende expertbeoordelaars zouden in een soort comité kunnen overleggen tot een overeenkomst wordt bereikt. Een praktisch alternatief zou bestaan uit het beoordelen van de gegevens (verzameld via een of andere alternatieve evaluatievorm) op basis van een aantal leidvragen die refereren naar relevante dimensies in de onderzochte inhoud en naar prestatiecriteria. Op die manier kan men betrouwbaarheid conceptualiseren als de bevestiging van de bevindingen van een eerste beoordelaar door een tweede. De tweede beoordelaar moet aangeven in welke mate de eerste beoordeling naar zijn mening al dan niet consistent is met de beschikbare gegevens.

Een eerder pragmatisch kwaliteitscriterium betreft de efficiëntie van een evaluatie. Op dit punt scoren de alternatieve evaluatievormen slechter dan vele traditionele evaluatiemethoden. Het implementeren en gebruiken van alternatieve evaluatiemethoden is kostelijk, tijdrovend en arbeidsintensief. Maeroff (1991, p.53) schrijft in dit opzicht: 'Expenses and time may turn out to be the brakes on the alternative assessment movement'. Te denken valt hierbij aan het ontwerpen van authentieke evaluatietaken en aan het opstellen van criteria en standaarden om die taken te evalueren. Verder is ook het ontwikkelen van criteria en standaarden zeer tijdrovend, terwijl deze activiteiten daarenboven een zekere expertise vereisen. Het afnemen van de evaluatie is zelfs voor experts in alternatieve evaluatievormen zeer tijdrovend. Dit heeft in belangrijke mate te maken met de aard van de te evalueren gegevens (essays, papers, portfolio's) en de manieren waarop vaak bijkomende informatie ingewonnen wordt (interviews, groepsdiscussies en dergelijke).

Tenslotte richten we ons op de rechtvaardigheid van alternatieve evaluatievormen. Omwille van het authentieke karakter van alternatieve evaluatiemethoden lijkt het alsof ze rechtvaardiger zijn dan traditionele testen. Onder rechtvaardiger wordt hier verstaan dat de beoordeling beter aansluit bij de reële prestaties van bepaalde minderheidsgroepen dan de traditionele testen. Empirische studies hieromtrent leiden echter niet tot een eenduidige conclusie. Indien er toch onrechtvaardige verschillen tussen bepaalde leerlinggroepen bestaan, dan moeten deze vaak worden toegeschreven aan het feit dat niet alle leerlingen op een even adequate manier werden voorbereid op de alternatieve evaluatiemethoden (Birenbaum, 1995). Bij het implementeren van alternatieve evaluatievormen moeten dan ook substantiële veranderingen worden doorgevoerd in de onderwijsstrategieën en -praktijk zelf. Bijvoorbeeld: de inhoud van de assessment moet alle leerlingen aanspreken en moet aansluiten bij de doelen die werden vooropgesteld; vertekende elementen (zoals taalgebruik) moeten zoveel mogelijk vermeden worden; en men moet een zekere openheid in stand houden omtrent het ontwikkelen, afnemen, scoren en rapporteren van de evaluatieresultaten.

Samenvatting

Omdat op basis van de resultaten van een evaluatie vaak belangrijke beslissingen worden genomen, is het noodzakelijk dat voldoende aandacht wordt besteed aan verschillende kwaliteitseisen. Deze eisen kunnen garanderen dat men correcte uitspraken doet op basis van de evaluatiegegevens. Hiertoe moeten gegevens op een doelmatige manier verzameld worden: dit betekent dat er over moet gewaakt worden dat er op een correcte (of betrouwbare) manier de juiste (of valide) gegevens worden verzameld en dit op een zo efficiënt mogelijke manier. Daarnaast is het belangrijk dat de evaluatie ook voldoende billijk (of rechtvaardig) is. Je kan spreken van een rechtvaardige evaluatie indien er sprake is van objectiviteit, doorzichtigheid en normering. Elk van deze kwaliteitseisen werd hoger toegelicht. Hierbij vestigden we de aandacht op twee bemerkingen, namelijk het variabel belang van de diverse kwaliteitseisen en de spanning tussen de kwaliteitseisen. De eerste bemerking houdt in dat naargelang van de functie van de evaluatie en de aard van de beslissingen die men op basis van de evaluatiegegevens wil nemen, de verschillende kwaliteitseisen in meer of mindere mate belangrijk zijn. Met de nadruk op de spanning tussen de kwaliteitseisen willen we de lezer ervan bewust maken dat bijvoorbeeld een hoge validiteit vaak ten koste gaat van de betrouwbaarheid en de objectiviteit van de evaluatie. Het tegemoetkomen aan de verschillende kwaliteitseisen is dan ook geen kwestie van alles of niets. Het is daarom raadzaam de kwaliteitseisen te benaderen als gradaties van meer of minder validiteit, betrouwbaarheid, doorzichtigheid terwijl steeds getracht wordt zoveel mogelijk aan alle eisen tegemoet te komen.

Naast de doelmatigheid en billijkheid als algemene kwaliteitskenmerken beschreven we in dit hoofdstuk de centrale kenmerken van alternatieve evaluatievormen. Alternatieve evaluatievormen kunnen echter niet zomaar zwart-wit worden afgezet tegenover andere evaluatiemethoden. De term 'alternatieve evaluatievormen' is bovendien een verzamelnaam. Dit neemt niet weg dat er specifieke kenmerken naar voren werden geschoven. Naargelang deze verschillende kenmer-

ken al dan niet vervuld zijn, is er sprake van verschillende evaluatiemethoden. In deel 4 zullen deze verschillende evaluatievormen besproken worden. De kenmerken van alternatieve evaluatievormen die aan bod kwamen zijn de aangepastheid aan de kenmerken van de leerlingen, de integratie van instructie en evaluatie, de betrokkenheid van leerkrachten en leerlingen en het constructiegerichte en levensechte karakter van de gebruikte evaluatietaken.

We stelden bij de aanvang dat het de verwachting is dat alternatieve evaluatievormen in vergelijking met traditionele evaluatie-instrumenten beter geschikt zijn om significante en blijvende leeruitkomsten aan het licht te brengen. Hieruit zou men kunnen afleiden dat het gebruik van alternatieve evaluatievormen ook de vervulling van algemene kwaliteitseisen impliceert. Dit is echter slechts gedeeltelijk correct. Het gebruik van authentieke vaardigheidstaken kan bijvoorbeeld garanties bieden om tot valide evaluatiebeoordelingen te komen. Anderzijds is het zo dat bepaalde kenmerken van alternatieve evaluatievormen wel degelijk kunnen impliceren dat er aan andere kwaliteitseisen moet worden ingebonden. Een nadeel van het werken met authentieke vaardigheidstaken is bijvoorbeeld dat de afname van vaardigheidstaken (of de uitvoering ervan) langer duurt dan traditionele kennistoetsen en dat de beoordeling ervan veel moeilijker is. Dat betekent dat de efficiëntie en de objectiviteit in het gedrang kunnen komen, vanwaar het belang om zorgvuldig het evaluatieproces zoals in deel 3 voorgesteld te doorlopen.

Met deze deels wel en deels niet is echter niet alles gezegd. Het doel van het benadrukken van alternatieve evaluatievormen gaat echter verder dan het realiseren van bepaalde kwaliteitskenmerken. Alternatieve evaluatievormen stimuleren bijvoorbeeld dat op een ruimere manier gebruik wordt gemaakt van evaluatiegegevens en dat leerlingen op een actieve manier in het evaluatieproces betrokken worden. Ze vormen met andere woorden werkwijzen om de kanteling naar een assessmentcultuur te maken. Indien zorgvuldig aandacht wordt besteed aan het evaluatieproces en aan de algemene kwaliteitskenmerken is de meerwaarde van alternatieve evaluatievormen dus duidelijk groter dan deze van een goede evaluatie zonder meer. Bij het presenteren van een aantal alternatieve evaluatievormen zullen we die stelling beargumenteren en illustreren.

Deel

3

Aandacht voor het evaluatie- proces

INHOUD - DEEL 3

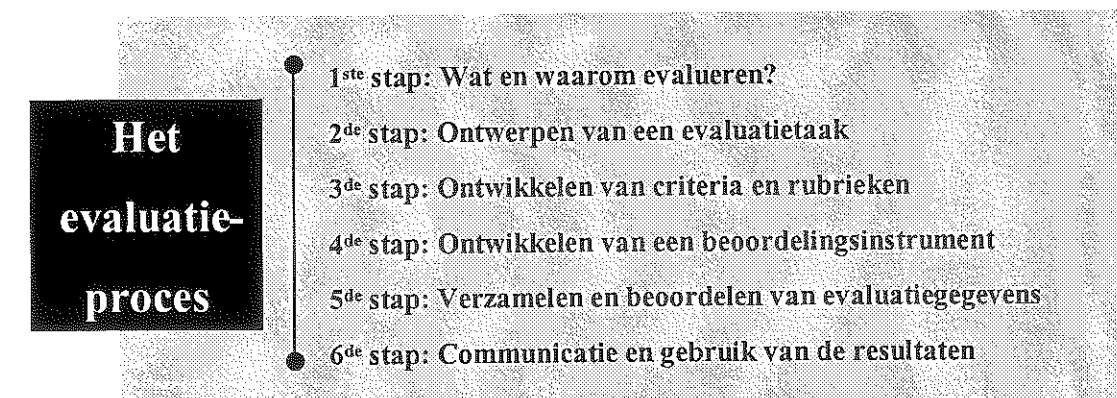
Aandacht voor het evaluatieproces

Inleidend	113
1 Stap 1: Wat en waarom evalueren?	114
1.1 De te evalueren doelen	114
1.2 Doelen/functionaliteiten van de evaluatie	116
1.3 Publiek van de evaluatie	117
2 Stap 2: Ontwerpen van een evaluatietask	118
2.1 Een overzicht van mogelijke evaluatietaken	119
2.2 Eisen aan specifieke evaluatietaken	123
2.3 Het ontwikkelen van geschikte evaluatietaken	124
2.4 Het beschrijven van een evaluatietask	125
3 Stap 3: Het ontwikkelen van criteria en rubrieken	126
3.1 Criteria, standaarden en verwachtingen?	127
3.2 Het belang van criteria en standaarden	130
3.3 Hoe kan je criteria en rubrieken ontwikkelen?	133
3.4 Het beoordelen van criteria en rubrieken	145
4 stap 4: Ontwikkelen van een beoordelingsinstrument	149
4.1 Het beoordelen van evaluatiegegevens	149
4.2 Over beoordelingsschalen	154
5 Stap 5: Verzamen en beoordelen van evaluatiegegevens	161
5.1 Wanneer spreekt men van een kwaliteitsvol beoordelingsproces?	162
5.2 Hoe een betrouwbaar waarderingsproces realiseren?	162
6 Stap 6: Communicatie en gebruik van de resultaten	165
6.1 Scoringsstrategieën	165
6.2 De communicatie van evaluatieresultaten	167
7. Opmerkingen bij het voorgestelde evaluatieproces	170

Inleidend

Om een doelmatige en billijke evaluatie te realiseren, moet zorgvuldig aandacht worden besteed aan het ontwikkelingsproces van de evaluatie. In het vorige deel werd dat herhaaldelijk als een belangrijke garantie voor kwaliteitsvolle evaluatie naar voren geschoven. Dat roept uiteraard de vraag op naar de manier waarop een goede (alternatieve) evaluatie vorm kan krijgen. Dit derde deel van het handboek wil -door aandacht te besteden aan het evaluatieproces- deze vraag beantwoorden.

Wanneer je als leerkracht een evaluatie op het getouw wil zetten, moeten een aantal stappen doorlopen worden. Deze stappen lichten we in dit deel toe. De structuur die aan de grondslag ligt van deze beschrijving bestaat in het doorlopen van het hele evaluatieproces. In het verdere verloop van dit deel staan we uitvoerig stil bij het evaluatieproces als dusdanig en illustreren we zowel dit proces als een aantal kritische succesfactoren. We willen benadrukken dat het niet onze bedoeling is een vastliggend stramien te presenteren. Het is immers noodzakelijk het evaluatieproces met de nodige flexibiliteit en creativiteit te benaderen. Wanneer je leerkrachten wil illustreren hoe een goede evaluatie ontwikkeld kan worden, dan kan je twee wegen bewandelen: of je structureert het verhaal op basis van alle kenmerken waaraan een evaluatie dient te voldoen, of je structureert het naar de chronologie van het ontwikkelingsproces. Het eerste zou eigenlijk de beste weg zijn; er is immers geen vastliggend recept of een strikte werkwijze om evaluaties te ontwikkelen, zolang er maar aan de noodzakelijke kwaliteitseisen tegemoet wordt gekomen. Toch hanteren wij in dit deel de chronologische benadering. We willen de lezer immers een houvast bieden door één mogelijke werkwijze uit de doeken te doen die garanties biedt voor het opzetten van kwaliteitsvolle evaluaties. Op die manier kan je als lezer zelf aan de slag, goed wetende dat het mogelijk is een andere invulling te geven aan het evaluatieproces.



Indien het evaluatieproces volledig doorlopen wordt, kunnen een zestal stappen onderscheiden worden. Elk van deze stappen komt in de volgende paragrafen aan bod.

Figuur 3.1: Structuur van het evaluatieproces

1 Stap 1: Wat en waarom evalueren?

Evalueren doe je niet zomaar. Bij de aanvang moet duidelijk worden gemaakt wat je wil bereiken met het opzetten van een evaluatie. De eerste stap in het selecteren en ontwerpen van een evaluatie bestaat dan ook uit het nagaan van de doelen van de evaluatie. Hierbij zijn drie vragen van belang (Herman et al, 1992; Mc Tighe & Ferrara, 1996):

1. Over welke aspecten van de leerlingprestaties (domeinen) wil je gegevens verzamelen? Wat wil je evalueren?
2. Waarvoor wil je de verkregen gegevens gebruiken? Wat is het doel van de evaluatie?
3. Voor wie zijn de evaluatiegegevens bedoeld? Met welk publiek voor ogen worden de evaluatiegegevens verzameld?

Deze vragen moeten met de nodige nauwgezetheid beantwoord worden en ze verdienen dan ook allemaal onze aandacht. Niet voor niets wordt het begrijpen van het evaluatieprobleem - en concreter het beantwoorden van bovenstaande vragen - beschouwd als de eerste maar ook meest cruciale stap in het evaluatieproces (Nevo, 1995). In de volgende paragrafen gaan we dieper in op zowel de te evalueren doelen als op de doelen en het publiek van de evaluatie op zich.

1.1 De te evalueren doelen

Het opzetten van een goede evaluatie vereist dat men de doelstellingen, die men bij leerlingen wil bereiken, kent en duidelijk kan uitdrukken. De vragen binnen volgende denkoefening kunnen helpen de doelstellingen te definiëren.

Even stilstaan... 3.1: Inventariseren van doelen

Staat u alvorens te evalueren wel eens stil bij volgende vragen? Neem even een concrete evaluatie voor de geest en achterhaal wat de eigenlijke inhoud is van wat u wil evalueren.

- ✓ Wat verwacht u dat leerlingen kennen en kunnen na een bepaalde tijdsperiode (les, cursus, schooljaar)?
- ✓ Waartoe wilt u komen met het onderwijsleerproces? Wat wilt u bereiken?
- ✓ Welke ontwikkelingsdomeinen wilt u bij de leerlingen beïnvloeden?
- ✓ Welke (meta)cognitieve vaardigheden wilt u ontwikkelen?
- ✓ Welke sociale en affectieve vaardigheden moeten uw leerlingen ontwikkelen?
- ✓ Welke probleemtipes wil u dat uw leerlingen kunnen oplossen?

De antwoorden op deze vragen vormen de basis voor het opzetten van onderwijsactiviteiten en voor het bepalen van prioriteiten in de klas. Daarnaast zijn de onderwijsdoelstellingen uiteraard eveneens bepalend om uit te maken welke leerlingprestaties geëvalueerd zullen worden. De overeenkomst tussen de doelstellingen, het curriculum en de inhoud van de evaluatie is van

cruciaal belang. Zij waarborgt immers dat leerlingen de te evalueren kennis en vaardigheden effectief kunnen verwerven en dat de evaluatie bestaat uit betekenisvolle taken die een beroep doen op de meest belangrijk geachte kennis, vaardigheden en attitudes.

De te evalueren doelen kunnen aan de hand van de hierboven geschetste taxonomie overzichtelijk worden samengebracht (zie deel 2, hoofdstuk 1). Merk op dat we met deze eerste stap in het evaluatieproces niet zozeer willen benadrukken dat elke leerkracht eigen doelen moet gaan formuleren en inventariseren, maar wel dat men bij het vormgeven aan onderwijs én evaluatie steeds de vooropgestelde doelen centraal moet stellen. Naast het eigen oordeel van de leerkracht zijn er trouwens nog bijkomende informatiebronnen beschikbaar om bovenstaande vragen te beantwoorden. Zo kunnen bijvoorbeeld ook de eindtermen en leerplannen als ondersteuning gebruikt worden bij het expliciteren van de doelstellingen die men wenst te evalueren. Volgende activiteit kan aangewend worden om samen met collega's duidelijkheid te scheppen over belangrijke onderwijsdoelen.

Activiteit 3.1: Identificeren van droomdoelen

Of de eigenlijke onderwijsdoelen nu vastliggen in eindtermen dan wel ingegeven worden door persoonlijke motieven, het identificeren van droomdoelen maakt dat je als leerkracht een duidelijke gerichtheid en motivatie creëert. Ideaal lijkt de situatie waarin je met collega's nadenkt over droomdoelen en hierin een gemeenschappelijk streven kan vinden. Volgende activiteit is hiertoe een mogelijke werkwijze (gebaseerd op Rolheiser et al., 2000).

Vorm groepjes van drie leerkrachten uit een zelfde vakgebied. Elke leerkracht krijgt afwisselend een andere functie, namelijk interviewer, respondent en verslaggever. De interviewer stelt (een aantal van) de vragen uit de vorige denkoefening aan de orde (zie even stilstaan...1). De respondent mag op die vragen enigszins dromerig antwoorden. Na een aantal minuten worden de rollen omgewisseld. Binnen elk groepje tracht men te komen tot een inventaris van droomdoelen die eventueel aan de volledige groep kan worden voorgesteld. Reflectie en discussie staan hierbij centraal. Ook de weerslag van de droomdoelen op instructie- en evaluatiemethoden kan aan bod komen.

Met volgend antwoord van een groep leerkrachten werkzaam in de derde graad Verzorging (BSO) willen we de lezer stimuleren even te reflecteren over de eigen droomdoelen.

Mijn droomdoel is dat mijn leerlingen zich zelf verantwoordelijk zouden voelen voor hun leren. Dat ze niet denken dat ze alleen op school kunnen bijleren over bijvoorbeeld het omgaan met bejaarden, maar dat ze dus ook in het dagelijks leven leren aandachtig te zijn voor potentiële leerervaringen. Met dagelijks leven bedoel ik ook dat ze het belang en de relevantie van hun eigen ideeën, gevoelens en vragen leren onderkennen. Ik wil ze leren om op basis daarvan zelf hun leren in handen te nemen en te organiseren. Zelf moet ik er daartoe voor zorgen te begrijpen wat ze belangrijk vinden om op die manier interessante, betekenisvolle en uitdagende leerervaringen aan te bieden.

Afsluitend kunnen we het vorige -hoe vanzelfsprekend ook- als volgt samenvatten: om een kwaliteitsvolle evaluatie vorm te geven, moeten de nagestreefde leerdoelen centraal gesteld worden.

1.2 Doelen/functies van de evaluatie

Naast de doelstellingen die men wenst te evalueren, moeten ook de doelstellingen van de evaluatie zelf geëxpliciteerd worden. Er is dus een duidelijk onderscheid tussen de onderwijsdoelen die een evaluatie evalueert en de doelen van die evaluatie (in de betekenis van welke de functies ervan zijn). Ook de evaluatie kan op zich verscheidene doelen hebben. Deze doelen/functies van de evaluatie variëren naargelang het gebruik dat van de verkregen evaluatiegegevens zal gemaakt worden. Verschillende doelen vereisen telkens een andere evaluatiestrategie.

In deel werd een onderscheid gemaakt tussen de summatieve en formatieve functies van evaluatie. Beide groepen functies vereisen een eigen invulling van de kwaliteitseisen en de vorm van de op te zetten evaluatie. Het onderscheid tussen formatieve en summatieve evaluatiefuncties is echter niet de enige mogelijkheid om een onderverdeling te maken. De functies van evaluatie kunnen ook nog aan de hand van een andere invalshoek onderverdeeld worden, namelijk op basis van het onderscheid tussen leerling- en instructiegerichte evaluatie (Herman et al, 1992; Janssens, 2000). Naargelang de evaluatie tot doel heeft informatie te verzamelen ten behoeve van de leerlingen (het leerproces) dan wel van de leerkracht (het instructieproces) is de functie ervan onder te brengen in één van deze twee groepen. Dit neemt niet weg dat ook combinaties binnen en tussen beide groepen van functies mogelijk zijn. We overlopen hieronder de verschillende functies van evaluatie volgens deze invalshoek. Hiermee willen we aantonen dat de bijbehorende evaluatie telkens anders zal moeten worden ingevuld.

1. Leerlinggerichte functies van evaluatie

Binnen de leerlinggerichte functies van evaluatie onderscheiden we drie groepen. Wanneer het doel bestaat uit het **evalueren van de verwezenlijkingen** van leerlingen zal men de evaluatiegegevens meestal voor volgende doeleinden gebruiken: cijfers geven, oriëntatie, meten van vooruitgang, selectie, toekennen van een getuigschrift, plaatsen van leerlingen in een niveaugroep. Omdat de nadruk ligt op het nagaan of leerlingen bepaalde kennis en vaardigheden beheersen, richt men zich voornamelijk op de uitkomsten of het product van het leren. Er worden bovendien vaak gewichtige beslissingen genomen zodat maximale aandacht moet worden geschonken aan de in deel 2 gepresenteerde kwaliteitseisen. Het doel van de evaluatie kan evenwel ook bestaan uit **diagnose en verbetering**. In dat geval moet men de evaluatie richten op zowel het proces als het product van het leren. Mogelijke doelen die men met een dergelijke evaluatie wenst te realiseren, zijn een sterkte- en zwakteanalyse van leerlingen en nagaan welke aspecten leerlingen al dan niet goed beheersen. De evaluatie moet dan ook zoveel mogelijk feedback opleveren die leerlingen een duidelijk beeld geeft van hun capaciteiten en hen stuurt bij het opzetten van handelingsplannen. Een derde belangrijk doel van leerlinggerichte evaluatie bestaat uit het aanzetten van leerlingen tot reflectie over hun eigen leerproces en -product.

2. Instructiegerichte functies van evaluatie

Evaluatiegegevens kunnen op verschillende manieren worden aangewend om richting te geven aan de instructie. Bij instructiegerichte evaluatie zijn normaliter **desturing van het onderwijsleerproces** en het **aanpassen van het onderwijsleerproces** aan het niveau van de leerlingen de centrale functies. Het resultaat van soortgelijke evaluatie kan bijvoorbeeld bestaan uit het herhalen van een stuk leerstof, het differentiëren in niveaugroepen, het gebruiken van een andere werkvorm of het verhogen van het leertempo. De evaluatie moet de leerkracht informeren en niet zozeer de leerlingen. Bij deze functie van evaluatie wordt de aandacht voor het oplossingsproces sterk benadrukt. Enkel productevaluatie zou immers niet veel informatie genereren. Een andere (in wezen oneigenlijke doch vaak voorkomende) functie van instructiegerichte evaluatie is ten slotte het **handhaven van orde**. Soms is alleen al de dreiging met evaluatie voldoende om een ordeprobleem op te lossen. De leerkracht beschikt met evaluatie over een krachtig machtsmiddel. Hoewel soms enigszins te begrijpen, kan een dergelijk onecht gebruik van evaluatie vanuit didactisch oogpunt niet goedgekeurd worden (Janssens et al., 2000). We laten deze instructiegerichte functie van evaluatie dan ook verder buiten beschouwing.

Even stilstaan... 3.2: De functies van evaluatie in uw onderwijspraktijk

- ✓ Welke van deze functies vervult evaluatie in uw onderwijspraktijk?
- ✓ Heeft evaluatie telkens dezelfde functie of is er eerder sprake van een combinatie van functies?
- ✓ Staat u voor u een evaluatie ontwerpt altijd stil bij de functie van de evaluatie?
- ✓ Varieert de evaluatievorm die u hanteert naargelang de functie van de evaluatie?

De aard van de doelstellingen en de eventuele combinatie van evaluatiefuncties die men met de evaluatie wil bereiken, zijn beslissend voor het verdere verloop van het evaluatieproces. Naargelang het belang en het gewicht van de beslissing die men op basis van de evaluatiegegevens wil nemen, zullen de volgende stappen immers anders worden ingevuld. De functie van de evaluatie moet dus duidelijk geëxpliciteerd worden alvorens het eigenlijke evaluatieproces vorm kan krijgen. Onthoud hieruit ook dat er niet één specifieke evaluatievorm bestaat die in alle gevallen aangewezen is.

1.3 Publiek van de evaluatie

Ook het publiek waarvoor de evaluatiegegevens verzameld worden, is van belang bij de verdere planning van het evaluatieproces. Niet alleen de keuze van de evaluatiemethode wordt er door beïnvloed, ook de manier waarop de evaluatiegegevens later gecommuniceerd zullen worden. Traditioneel waren leerkrachten de enige betrokkenen bij het beoordelen van leerlingen. Tegenwoordig spelen ook andere betrokkenen een actieve rol in de evaluatie. Deze rol kan zich zowel toespitsen op het verzamelen van evaluatiegegevens, als op het beoordelen en terugkoppelen ervan. Met de vraag naar het publiek van een evaluatie wordt duidelijkheid geschept over voor wie de evaluatiegegevens bedoeld zijn. Dat het publiek van een evaluatie mogelijk zeer verscheiden is, blijkt uit volgende lijst van mogelijke doelgroepen: leerkracht(en), de leerling zelf, andere leerlingen, ouders, instellingen voor hoger onderwijs, de eigen school, centra voor

leerlingbegeleiding, toekomstige werkgevers, de overheid etc. Later zullen we aantonen dat de rol van derden zich niet hoeft te beperken tot die van geïnteresseerden in de evaluatieresultaten. Ook in de eigenlijke evaluatie kunnen zij een meerwaarde bieden. We denken dan aan de verschillende en leerrijke perspectieven die ze tijdens de evaluatie kunnen aanreiken en aan de waardevolle feedback die ze kunnen verzorgen.

In figuur 3.2 wordt een overzicht gegeven van de verschillende vragen die beantwoord moeten worden bij de eerste stap van het evaluatieproces. Een belangrijke les hierbij is dat, naast de boven beschreven verschillen in functies van de evaluatie, ook de leeruitkomsten die men wenst te evalueren een rol spelen bij de keuze van de meest gepaste (alternatieve) evaluatiemethode. Het expliciteren van het wat en het waarom van de evaluatie is dan ook essentieel bij de aanvang van een evaluatie (Fisher & King, 1995). Daarnaast moet vooraf gereflecteerd worden over de doelgroep waaruit het publiek van de evaluatie zal bestaan. Pas dan kan men zich de vraag stellen op welke manier de prestaties van leerlingen best geëvalueerd kunnen worden.

Figuur 3.2: Vragen bij het plannen van een evaluatie

Leeruitkomsten	Doelen van de evaluatie	Publiek
Wat moeten de leerlingen verstaan en kunnen?	Waarom evalueren we en hoe gaan we de evaluatiegegevens gebruiken?	Voor wie zijn de evaluatiegegevens bedoeld?
Bron: Conley, 1996.	<input type="checkbox"/> Diagnosticeren <input type="checkbox"/> Feedback geven <input type="checkbox"/> Optimaliseren van de instructie <input type="checkbox"/> Leerverwachtingen communiceren <input type="checkbox"/> Motiveren, aandacht en inspanningen van leerlingen stimuleren <input type="checkbox"/> Leerlingevaluatie (bv. attestering) <input type="checkbox"/> Evaluatie van de opleiding <input type="checkbox"/> ...	<input type="checkbox"/> Leerkracht <input type="checkbox"/> Leerlingen <input type="checkbox"/> Ouders <input type="checkbox"/> Hoger onderwijs <input type="checkbox"/> Bedrijfswereld <input type="checkbox"/> Schoolbeleid <input type="checkbox"/> CLB <input type="checkbox"/> Derden <input type="checkbox"/> ...

2 Stap 2: Ontwerpen van een evaluatietaak

Na het verwezenlijken van stap 1 beschikt men over een overzicht van de resultaten die men van leerlingen verwacht (kennis, inzicht, vaardigheden, attitudes). Essentieel voor een goede evaluatie is deze elementen te weerspiegelen in de eigenlijke evaluatietaken. De volgende stap in het evaluatieproces bestaat dan ook uit het selecteren en/of ontwerpen van taken die de bedoelde leerlingprestaties aan het licht kunnen brengen.

De waarde van een evaluatie hangt af van de mogelijkheden van de leerkracht om gepaste taken te vinden die leerlingen in staat stellen hun kennis en vaardigheden te demonstreren. Vandaar het belang van het selecteren van evaluatietaken die de leerlingen uitnodigen betekenisvolle en realistische problemen op te lossen. Leerlingen laten zich immers eerder van hun beste kant zien indien de evaluatietaak hen uitdaagt en interesseert (Fisher & King, 1995). Evaluatietaken

moeten bovendien tegemoet komen aan de doelstellingen en de inhoud die men in stap 1 vooropstelde. Dat betekent dat er geen sprake kan zijn van een over de hele lijn 'beste' of meest waardevolle soort evaluatietaak. Men moet zich bovendien steeds de vraag stellen of een andere evaluatiemethode niet beter geschikt is om de vooropgestelde gegevens te verzamelen. De informatie die men op basis van een bepaalde evaluatietaak of -opdracht verzamelt, moeten daarom steeds worden bekeken in het licht van de leerdoelen.

Ook bij deze tweede stap in het evaluatieproces verdienen een aantal elementen onze aandacht. In de volgende paragrafen besteden we bijkomende aandacht aan de eisen waaraan specifieke evaluatietaken dienen te voldoen en aan het ontwikkelen en beschrijven van evaluatietaken. Eerst schetsen we evenwel een overzicht van mogelijke evaluatietaken.

2.1 Een overzicht van mogelijke evaluatietaken

Evaluatiemethoden kunnen op verschillende manieren gecategoriseerd worden. De indeling die wij hier presenteren, maakt een onderscheid tussen vijf (deels overlappende) groepen: gesloten vragen, open vragen, evaluatie van producten, observatie van prestaties en evaluatie van het leerproces. Hierna beschrijven en illustreren we deze groepen. Terloops: bij het overlopen van de verschillende evaluatietaken kan het leerrijk zijn even stil te staan bij de vraag welke evaluatietaken u zelf gebruikt.

1. Gesloten vragen - te selecteren antwoorden

Bij de zogenaamde gesloten vragen bestaat de evaluatietaak erin dat leerlingen een antwoord selecteren uit een aantal vooraf ter beschikking gestelde mogelijkheden. De leerlingen moeten dus zelf geen antwoord construeren (Conley, 1996; Gredler, 1999). Over deze groep - die eerder kenmerkend is voor de traditionele evaluatiebenadering - kunnen we hier kort zijn. Merk op dat we hiermee niet willen uitsluiten dat deze groep van evaluatiemethoden voor het evalueren van bepaalde doelstellingen functioneel kan zijn. Voorbeelden van evaluatievormen die in deze groep thuishoren zijn:

- ✓ meerkeuzevragen;
- ✓ waar-vals vragen;
- ✓ invuloefeningen met te selecteren antwoorden;
- ✓ rangschikkings- en classificatievragen;
- ✓ verbindingsoefeningen.

2. Open vragen - te construeren antwoorden

Evaluatietaken die worden ondergebracht in de groep 'te construeren antwoorden' verwachten van leerlingen dat ze antwoorden produceren op open vragen en problemen. Ondanks het feit dat men meestal een zogenaamd correct en te verwachten antwoord beoogt (bijvoorbeeld bij

invuloefeningen) is het mogelijk dat een variëteit aan antwoorden wordt gegeven (in tegenstelling tot de groep 'te selecteren antwoorden'). Naargelang de aard van het antwoord dat men van leerlingen verwacht en de mate waarin leerlingen antwoordvrijheid krijgen, worden binnen de open vragen drie groepen onderscheiden. Het onderscheid tussen deze groepen is echter niet eenduidig te maken is.

a. Korte-antwoord-vragen

Korte-antwoord-vragen verwachten van leerlingen dat ze een antwoord formuleren dat bestaat uit een getal of een aantal woorden (een naam, datum, term of iets dergelijks). De vragen zijn meestal kort of bestaan uit korte uitspraken die moeten worden aangevuld. Dit vraagtype is in wezen enkel geschikt om op een valide manier geheugenkennis te evalueren (feiten, procedures).

Voorbeelden: Wat is de formule voor het berekenen van de omtrek van een cirkel? Welke zijn de drie laatste presidenten van Amerika? Met welke stijlfiguur benoemen we een door klanknabootsing gevormd woord?

b. Begrenzende-open-vragen

Bij begrenzendende-open-vragen is het antwoord dat men van leerlingen verwacht ruimer dan één enkel getal of woord, maar wel begrensd. Deze begrenzing kan zowel betrekking hebben op de lengte, de vorm als de inhoud van het antwoord (Janssens et al., 2000). Het antwoord van de leerlingen is enigszins te voorspellen in die zin dat de leerkracht een duidelijk zicht heeft op hoe het antwoord van de leerling er zal uitzien. Begrenzende-open-vragen kunnen vooral bij het evalueren van declaratieve en procedurele kennis doelgericht aangewend worden.

Voorbeelden van toetsvragen die in deze groep thuishoren zijn: Benoem de verschillende delen in volgend elektrisch schema. Omschrijf waarom metalen elektrisch geleidend zijn. Kippensoep is gedeeltelijk een oplossing, een suspensie en een emulsie. Leg uit.

c. Niet begrenzendende-open-vragen of essayvragen

Niet begrenzendende-open-vragen bieden leerlingen meer vrijheid om hun antwoord vorm te geven. Het is aan de leerling om een eigen antwoord te creëren met een eigen invulling van lengte, structuur en inhoud. Het antwoord op deze vragen bestaat veelal uit een essay. Dit is zoveel als een niet te korte verhandeling over een bepaald onderwerp. De leerling bepaalt hierbij zelf de gedachtegang, formuleert kerngedachten, bepaalt de indeling van de tekst (inleiding, midden, slot) en legt verbanden tussen verschillende elementen.

Voorbeelden: Wat is de invloed van massamedia op de manier waarop oorlogsconflicten beslecht worden? Zou de invulling van de rol van een schooldirecteur veranderen indien deze democratisch door de leerlingen verkozen zou worden? Welke rol spelen recente accentverschuivingen in onderwijsdoelen in de evolutie van een traditionele toetscultuur naar een assessmentcultuur?

3. Evaluatie van producten

De producten van leerlingen bieden tastbare indicaties van het toepassen van kennis en vaardigheden. De meerwaarde van de evaluatie van producten bestaat in het feit dat de evaluatie als het ware 'authentiek' is in die zin dat het proces sterk aansluit bij wat in levensechte situaties gebeurt. Productevaluaties tonen aan wat de leerling kan terwijl tevens aandachtspunten voor verbetering aan het licht worden gebracht. Volgende opsomming presenteert een greep uit de verschillende producten van leerlingen die de basis kunnen vormen om evaluatieve beoordelingen uit te spreken:

- ✓ Geschreven producten
 - essays over allerlei onderwerpen
 - tijdschriftartikelen, brieven, gedichten
 - onderzoeks- en laboverslagen
 - stageverslagen, logboeken
- ✓ Visuele en auditieve producten
 - multidimensionele modellen
 - tentoonstellingen
 - video's of tapes van een presentatie
- ✓ Werkstukken: handwerk, een kleerkast maken, lasverbindingen
- ✓ En verder de portfolio: een verzameling van verscheidene producten (zie deel 4).

Het onderscheid tussen niet begrenzendende-open-vragen en de evaluatie van producten is niet altijd even duidelijk te maken. Een essay kan bijvoorbeeld zowel opgevat worden als een product en als een antwoord op een niet begrenzendende-open-vraag. Om de verschillen tussen beide groepen evaluatietaken te duiden moeten we een duidelijk onderscheid maken tussen toets- en evaluatievragen enerzijds en opdrachten anderzijds. Het verschil situeert zich met andere woorden op het niveau van de evaluatietask die men leerlingen voorschotelt en niet op het niveau van het antwoord van de leerlingen. Tot welke groep van evaluatietaken een antwoord van een leerling in de vorm van een essay behoort, hangt dus af van de evaluatietask die eraan ten grondslag lag. *Even een voorbeeld. Neem het essay dat resulteerde uit volgende niet begrenzendende-open-vraag: hoe kan men een evaluatieproces best inrichten zodat de evaluatie tegemoetkomt aan doelmatigheids- en billijkheidseisen? In dit geval is er sprake van een antwoord. Indien hetzelfde essay het resultaat zou zijn van de volgende opdracht zou er echter sprake zijn van een product: Als domeinverantwoordelijke voor evaluatie binnen het leerkrachtenteam in jouw school krijg je de vraag op welke manier evaluatie best vorm kan krijgen zodat aan de doelmatigheids- en billijkheidseisen wordt tegemoetgekomen. In een bondig document werk je daartoe een model tot evolutieproces uit dat jouw collega-leerkrachten op weg helpt.*

4. Observatie van prestaties

Door de prestaties van leerlingen in de evaluatie centraal te stellen, kunnen leerkrachten direct de toepassing van kennis en vaardigheden door de leerling observeren. De directe aansluiting van de realistische, levensechte opdrachten bij de nagestreefde prestaties maakt dat deze groep

van evaluatiemethoden als het meest authentiek wordt bestempeld (Conley, 1996). Observaties tijdens stages of opdrachten in praktijklokalen benaderen wellicht het meest de werkelijkheid. Hierbij denken we dan bijvoorbeeld aan het wassen van een bejaarde en het aansluiten van een elektriciteitskast. Toch kunnen ook in het klaslokaal aan de hand van simulaties levensechte evaluatietaken georganiseerd worden. Een aantal voorbeelden zijn: het geven van een presentatie, het uitvoeren van een experiment, rollenspelen en het houden van een debat.

Net als producten zijn geobserveerde prestaties vaak moeilijker op een objectieve manier te beoordelen. Vandaar het belang van het zorgvuldig doorlopen van de volgende stappen in het evaluatieproces. We willen terloops benadrukken dat de rapportering van de geobserveerde gegevens naargelang de doelen en het belang van de evaluatie een andere invulling kan krijgen. De observator kan bijvoorbeeld vrij aantekeningen maken van het geobserveerde gedrag of hij kan gebruik maken van een checklist of een beoordelingsschaal. Beide vormen van rapportering komen later in stap 4 aan bod.

5. Evaluatie van het leerproces

Procesgerichte evaluatie beoogt informatie te verzamelen over leerstrategieën en denkprocessen van de leerlingen. Eerder dan de blik te richten op tastbare leeruitkomsten, leggen deze evaluatiemethodes de nadruk op cognitieve processen bij de leerlingen. Het doel is te voorzien in diagnostische informatie als basis voor feedback. Ook voor het ontwikkelen van metacognitieve vaardigheden worden deze werkvormen gehanteerd (Conley, 1996).

De methodes om het leerproces te evalueren, zijn verscheiden. Sommige van volgende voorbeelden worden door leerkrachten permanent gebruikt terwijl andere eerder uitzonderlijk zijn:

- ✓ mondelinge vragen, interviews;
- ✓ observaties;
- ✓ luidop denken;
- ✓ leerlogboeken.

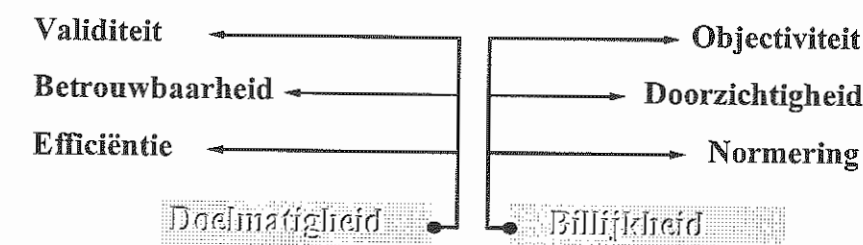
Uit bovenstaand overzicht blijkt duidelijk dat leerkrachten een verscheidenheid aan evaluatiemethodes ter beschikking hebben. De moeilijkheid bestaat erin bij elke evaluatie telkens een adequate keuze te maken en gepaste taken te ontwikkelen. Gezien de invalshoek van de zogenaamd alternatieve evaluatievormen beperken we ons in de verdere bespreking tot constructiegerichte en levensechte evaluatietaken. Bij het beschrijven van een aantal alternatieve evaluatievormen in deel 4 zullen we met andere woorden uitvoeriger aandacht besteden aan een aantal van de gepresenteerde evaluatietaken (zoals open toetsvragen, observatie van prestaties, authentieke taken, portfolio's en evaluatie van producten).

2.2 Eisen aan specifieke evaluatietaken

De evaluatietaak is van cruciaal belang bij het ontwerpen van een goede (alternatieve) evaluatie. In grote lijnen kunnen daarom voor de evaluatietaak dezelfde eisen worden vooropgesteld als

voor de evaluatie in haar geheel. Bij het kiezen of ontwerpen van evaluatietaken moet men dus tegemoet komen aan de eisen van goede evaluatie in het algemeen en van alternatieve evaluatievormen in het bijzonder (zie deel 2). Door de kenmerken van alternatieve evaluatievormen te formuleren als criteria bekwamen we als het ware een checklist die een leidraad kan zijn bij het ontwikkelen van evaluatietaken en die een kritisch licht kan werpen op reeds ontwikkelde taken. Om de eisen aan specifieke evaluatietaken te benadrukken, verwijzen we naar figuur 3.3 en naar de bijbehorende denkoefening.

Figuur 3.3: Kenmerken van goede evaluatie



Kenmerken goede alternatieve evaluatie

- ✓ Kennisconstructie is het uitgangspunt, niet kennisreproductie
- ✓ De evaluatie is aangepast aan de kenmerken van de leerlingen
- ✓ De integratie van evaluatie met onderwijzen en leren
- ✓ Betrokkenheid van leerkrachten en leerlingen
- ✓ Gebruik van authentieke of levensechte situaties
- ✓ Zowel het proces als het product van het leren zijn belangrijk

Even stilstaan... 3.3: Invulling van kwaliteitseisen aan evaluatietaken

In volgende alinea verduidelijken we wat bovenstaand schema voor de te hanteren evaluatietaken impliceert. Alvorens verder te lezen is het aangeraden deze denkoefening zelf eerst even te maken (bijkomende informatie over de verschillende aspecten in het schema vindt u in deel 2).

Samengevat klinken de eisen aan evaluatietaken als volgt. Een evaluatietaak moet zo gericht als mogelijk gegevens verzamelen over de leerdomeinen die je als leerkracht vooropstelde. Hierbij moeten eventuele toevalligheden en meetfouten tot een strikt minimum beperkt worden, en dit zonder de noodzakelijke tijdsinvestering te groot te maken. De taak moet voor alle leerlingen duidelijk zijn en ze moet een rechtvaardige en legitimeerbare beoordeling mogelijk maken. Leerlingen moeten aan de slag, ze moeten zelf een antwoord construeren en dit, indien mogelijk, in een zo levensechte context. Bij het beantwoorden van de vraag en bij het uitvoeren van de opdracht dragen ook de leerlingen verantwoordelijkheid over bepaalde beslissingen. Waar nodig wordt de taak gedifferentieerd om aan specifieke kenmerken van individuele leerlingen tegemoet te komen.

2.3 Het ontwikkelen van geschikte evaluatietaken

Essentieel in deze tweede stap van het evaluatieproces is de vraag op welke manier goede evaluatietaken ontworpen kunnen worden. Het komt er met andere woorden op neer leerlingen taken voor te leggen die hen in staat stellen de vooropgestelde kennis, vaardigheden en attitudes te demonstreren. We stelden hoger dat de kwaliteit van de in Vlaanderen sterk ingeburgerde begrenzende-open-vragen vaak te wensen overlaat (*zoals bijvoorbeeld geldt voor de toetsvraag: noem drie problemen die zouden kunnen ontstaan indien de aarde langzaam opwarmt*). De meest eenvoudige manier om de kwaliteit van dergelijke taken te verbeteren, bestaat in het vertrekken van en verder uitwerken van reeds bestaande (evaluatie)taken. Met wat creativiteit en inlevingsvermogen is de succesvolle omschakeling naar prestatietaken vaak minder veraf dan gedacht. Bijvoorbeeld: *Talrijke wetenschappers slaan alarm omdat de aarde dreigt op te warmen. De man in de straat ligt daar echter niet van wakker en is niet geneigd om zijn manier van leven aan te passen. Stel een pamflet (1 half A4) op waarin je aantoont op welke manieren de opwarming van de aarde ook de man in de straat aanbelangt. Vrijdagavond verspreiden we het pamflet aan de schoolpoort.*

Voor het ontwikkelen van uitdagende en levensechte vaardigheidstaken kunnen verschillende benaderingen aangewend worden. We zetten illustratief een drietal strategieën op een rij: de van-buitenuit-strategie, de van-binnenuit-strategie en achterwaarts plannen (ASCD, 1996).

a. De van-buitenuit-strategie

De van-buitenuit-strategie bestaat erin dat het verband van de te evalueren doelstellingen met het curriculum reeds duidelijk is, maar dat voor de specifieke doelstellingen nog een gepaste taak gevonden moet worden. Dit is in principe de normale gang van zaken: je hebt een doelstelling voor ogen en je zoekt hiervoor een geschikte evaluatietask. Door te brainstormen met collega's en/of leerlingen en door ideeën en reeds bestaande taken uit te wisselen, kunnen voor het vooropgestelde doel verscheidene taakvoorstellen verzameld worden. Vervolgens komt het er op aan uit deze lijst van taakvoorstellen die taak te selecteren die het best de voor de vooropgestelde doelstellingen noodzakelijke geachte kennis, vaardigheden en attitudes aan het licht kan brengen.

b. De van-binnenuit-strategie

Deze strategie bestaat erin te vertrekken van een activiteit waarvan je weet dat de leerlingen ze motiverend, aangenaam en uitdagend vinden. Dit kan dus best een reeds bestaande taak zijn. Vervolgens stel je jezelf de vraag of die taak niet verbonden kan worden aan andere inhouden, vaardigheden en attitudes die onderdeel uitmaken van het curriculum. Je weet met andere woorden wat leerlingen uitdagend vinden en je zoekt doelstellingen die je hierbij aansluitend kan evalueren. Indien leerlingen bijvoorbeeld graag in groep werken en het motiverend vinden om verscheidene ICT-toepassingen te gebruiken, kan je bij het creëren van evaluatietaken hier handig op inspelen. Als blijkt dat leerlingen sterker gemotiveerd zijn wanneer leertaken aansluiten bij hun eigen leefwereld, kan je daar bij het ontwerpen van evaluatietaken rekening mee houden.

c. Achterwaarts plannen

Een derde strategie bestaat erin te vertrekken van de eindtaak die leerlingen geacht worden aan het einde van de cursus of van het leerjaar succesvol te kunnen uitvoeren. Deze taak kan dan bijvoorbeeld bestaan uit een probleem dat professionals in hun dagelijkse beroepspraktijk moeten aanpakken. Vanuit die taak kan je vervolgens systematisch eenvoudigere taken opstellen tot een aan de leerlingen aangepast niveau bereikt wordt. Elk van die taken draagt in dat geval bij tot het succesvol uitvoeren van de eindtaak, is zelf uitdagend en is duidelijk aan het curriculum verbonden.

Welke weg ook gevolgd wordt voor het ontwerpen van een evaluatietask, het ontwikkelingswerk vergt enige oefening en creativiteit. Voornamelijk de combinatie van het uitdagende karakter en de doelgerichtheid van de beoogde evaluatietaken maken dat het creëren ervan de nodige aandacht verdient. Bij het bespreken van vaardigheidsevaluatie en authentieke evaluatie als vormen van alternatieve evaluatie staan we uitvoeriger stil bij de kenmerken, de types en het opstellen van alternatieve evaluatietaken (zie deel 4).

2.4 Het beschrijven van een evaluatietask

Ter afronding van de tweede stap van het evaluatieproces vestigen we de aandacht op een vorm van kwaliteitsbewaking die aansluit bij het creëren van evaluatietaken, namelijk het beschrijven ervan. De aard van de beslissingen die worden gebaseerd op de evaluatiegegevens maken het wenselijk dat de evaluatietaken voldoende gespecificeerd en gedocumenteerd worden. De verschillende aspecten van een evaluatietask moeten met andere woorden duidelijk omschreven worden. Op die manier kunnen anderen de beschikbare evaluatiegegevens interpreteren en eventueel de methode (c.q. de evaluatietask) herhalen bij andere en/of dezelfde leerlingen in andere en/of dezelfde onderwijssettings. Het beschrijven van evaluatietaken is dus van belang bij het garanderen van een betrouwbare evaluatie. Het is interessant om later de uitgebreide voorbereiding en beschrijving van een specifieke evaluatietask opnieuw ter hand te kunnen nemen. Het beschrijven van evaluatietaken is daarenboven eveneens aangewezen omwille van het feit dat het de leerkracht aanzet tot reflectie. Op die manier worden immers problemen of elementen die over het hoofd gezien zouden kunnen worden, toch vroegtijdig gedetecteerd. Bovendien wordt de ontwikkelaar of de gebruiker van de taak aangezet om na te gaan of de taak inderdaad aan de nodige kwaliteitseisen voldoet. We willen benadrukken dat de beschrijving van een evaluatietask geen louter administratieve bezigheid is. Evaluatietaken worden immers niet alleen beschreven om de leerkracht(en) te informeren, maar ook om voor de leerlingen duidelijkheid te scheppen omtrent wat er van hen verwacht wordt. Evaluatietaken in een assessmentcultuur zijn vaak uitgebreider en informatiever dan traditionele evaluatietaken. Door de taken te beschrijven en de informatie ook aan leerlingen te bezorgen, wordt de transparantie van de evaluatie vergroot.

De verschillende elementen die bij het beschrijven van een evaluatietask zouden moeten worden opgenomen, zijn afhankelijk van de aard van de betreffende evaluatie. Dit neemt niet weg

dat er toch een aantal algemene aspecten naar voren kunnen worden geschoven die het door- gaans waard zijn om op papier te worden vastgelegd. Aan de hand van de elementen en vragen in figuur 3.4 presenteren we een mogelijk sjabloon voor het beschrijven van een evaluatietaak.

Figuur 3.4: Ondersteunende vragen bij het beschrijven van evaluatietaken

Titel van de taak:		Leerdomein:
Ontwerper:	Studiejaar:	Datum:
1. Doelen van de taak. Welke leerlingprestaties wil de taak evalueren? 2. Waarvoor zullen de evaluatiegegevens gebruikt worden? Wie zal ze gebruiken? 3. Beschrijving van de evaluatietaak en -context: ✓ Wat is de precieze taak of opdracht die de leerlingen moeten oplossen? ✓ Hoe ziet de evaluatie er praktisch uit? Hoe ziet het tijdschema eruit en wat is de volgorde van de taken? Welke instructies worden aan de leerlingen gegeven? Moet de taak individueel of in groep worden uitgevoerd? Zijn er verschillende rollen? Welke hulpmiddelen krijgen leerlingen ter beschikking? Waar moet de taak opge- lost worden (thuis, op school)? ✓ Welke voorkennis moeten leerlingen hebben om de taak tot een goed einde te brengen? 4. Welke procedures zullen gehanteerd worden bij het waarden van evaluatiegege- vens? Welke criteria en rubrieken worden vooropgesteld? Wie zal evalueren? Welke schaal zal gebruikt worden om de evaluatiegegevens te beoordelen?		

Aan de hand van de selectie vragen in figuur 3.4 wordt het mogelijk de essentiële aspecten van de evaluatietaak te beschrijven. Een aantal hiervan komen pas in de latere stappen van het eva- luatieproces aan bod. Het beschrijven van een evaluatietaak moet daarom ruimer opgevat wor- den dan het beschrijven van de taak op zich. Zowel de context van de evaluatie als de evaluatie- criteria en de gebruikte schalen kunnen toegelicht worden om de beschrijving maximaal te laten renderen. Het beschrijven van evaluatietaken moet dan ook eerder bekeken worden als een ge- leidelijk iets waarbij gaandeweg in het evaluatieproces steeds meer informatie wordt opgeno- men. Evaluatiecriteria kunnen bijvoorbeeld niet beschreven worden voor ze ontwikkeld werden.

3 Stap 3: Het ontwikkelen van criteria en rubrieken

Zonder evaluatiecriteria zou een evaluatietaak in wezen niet verschillen van een leertaak of instructieactiviteit. De criteria die gehanteerd worden om de prestaties van leerlingen te beoor- delen, zijn dan ook een sleutelement in het opzetten van een evaluatie. Alternatieve evaluatie- vormen nodigen leerlingen uit een verscheidenheid aan mogelijke antwoorden te geven. Derge- lijke complexe antwoorden kunnen niet simpelweg als juist of fout beoordeeld worden (zoals bij antwoorden op meerkeuzevragen en korte-antwoord-vragen). Bovendien wordt bij alternatieve evaluatievormen vaak aandacht besteed aan het oplossingsproces. Om hierover oordelen uit te spreken, moet men beschikken over duidelijke beoordelingscriteria of scoringsrichtlijnen.

Pas wanneer het ontwikkelen en het gebruik van de criteria aan bepaalde voorwaarden voldoet, kunnen de criteria ook effectief bijdragen tot kwaliteitsvol evalueren. In volgende paragrafen gaan we dieper in op deze voorwaarden. We maken hierbij een onderscheid tussen het eigenlij- ke ontwikkelingsproces van evaluatiecriteria en een aantal kenmerken van goede criteria. Eerst staan we echter stil bij de definitie van en de noodzaak aan criteria, standaarden, verwachtingen en rubrieken.

3.1 Criteria, standaarden en verwachtingen?

Alvorens elk van deze termen kort te beschrijven, illustreren we aan de hand van volgende denkoefening waarom criteria, standaarden en verwachtingen een centrale rol spelen in het evaluatieproces.

Even stilstaan... 3.4: Het belang van evaluatiecriteria

Een belangrijk doel van het onderwijs is het ontwikkelen van de capaciteiten van leerlingen om in groep te werken. Laten we ervan uitgaan dat je als leerkracht tijdens een taak twee leerlingen observeerde (Michel en Dominique) en kort je vaststellingen neerschreef. Je hebt met andere woorden aan de hand van een evaluatietaak een aantal evaluatiegegevens verzameld. Lees deze gegevens eerst aandachtig door en geef daarna een antwoord op de onderstaande vragen.

Michel: Wat onmiddellijk opvalt, is dat Michel een goede, actieve luisteraar is. Hij zal meestal pas zelf argumenten aanbrengen wanneer hij vermoedt dat ze anders over het hoofd zouden worden gezien. In dat geval zal hij voorzichtig een inbreng proberen te doen, maar dan wel zon- der zichzelf te willen opdringen. Meer zelfs, indien de rest van de groep niet onmiddellijk aan- dacht besteedt aan Michels inbreng, dan volhardt hij niet. Zijn inbreng is dus niet altijd even groot, tenzij er inhoudelijke of relationele conflicten opduiken. Uit onverwachte hoek komt hij dan aanzetten met een compromis en benadrukt hij het belang van positieve relaties binnen de groep. Je zou kunnen zeggen dat Michel zich als groepslid goed aanpast aan de situatie in de groep. Wanneer de groep taken verdeeld heeft, dan kan je er steeds op rekenen dat Michel zijn individuele bijdrage goed voorbereid heeft.

Dominique: Dominique organiseert spontaan de praktische uitvoering van de taak. Ze sugge- reert potentiële werkwijzen en bewaakt dat aan bepaalde tijdslimieten wordt tegemoetgekomen. Ze neemt dan ook minstens gedeeltelijk het leiderschap binnen de groep op, ze stuurt en geeft richting. Hierbij moedigt ze anderen aan om op een gedreven manier mee te werken. Daarte- genover staat dat haar eigen inbreng niet altijd als even constructief ervaren wordt door de an- dere groepsleden. Wanneer er beslissingen moeten worden genomen, dan durft ze tegenargu- menten uit het oog verliezen en blijkt ze geen grootmeester in het sluiten van compromissen.

Hierbij aansluitend drie vragen:

1. Welke van beide leerlingprestaties vindt u het meest waardevol? Waarom?
2. Welke problemen ervaart u bij het beantwoorden van deze vraag?
3. Hoe zou u zelf tewerk gaan om de vaardigheden van deze individuele leerlingen om in groep te werken op een betrouwbare en objectieve manier te beoordelen?

Bij het beoordelen van de prestaties van leerlingen kan men niet voorbijgaan aan het gebruik van criteria, standaarden en verwachtingen. Dat geldt des te meer bij het beoordelen van leerlingprestaties die aan de hand van alternatieve evaluatievormen verzameld werden. Hierna beschrijven we eerst de betekenis van de drie begrippen, daarna besteden we aandacht aan het belang van criteria in het evaluatieproces (zie paragraaf 3.2).

3.1.1 Wat zijn criteria?

Criteria vormen de basis waarop we oordelen. Het zijn de voorwaarden waaraan een (leerling)prestatie dient te voldoen om als succesvol te worden bestempeld (Wiggins, 1996). In het dagelijks leven gebruiken we criteria veelal onbewust zonder ze -tenzij er naar gevraagd wordt- te expliciteren. Bij het beoordelen van een film bijvoorbeeld baseren we ons op criteria zoals de verhaallijn, de acteerpresetaties, het camerawerk en dergelijke. Ook al zijn criteria dus vaak niet neergeschreven en niet geformaliseerd, toch beïnvloeden ze onze beslissingen. In het onderwijs is dat niet anders. Bijvoorbeeld bij het beoordelen van wat al dan niet goede prestaties en goed gedrag zijn. De criteria geven aan welke de meest belangrijk geachte leeruitkomsten zijn en welke kennis en vaardigheden leerlingen moeten kunnen demonstreren. Criteria geven aan wat leerlingen moeten doen en welke aspecten van een prestatie in de beoordeling van tel zullen zijn. Figuur 3.5 illustreert het begrip criteria.

Figuur 3.5: Een selectie van criteria

Mogelijke criteria voor de leerlingprestatie 'spelen met kinderen' (BSO 3^{de} graad Verzorging):

- ✓ Vragen stellen om de belangstelling van het kind te achterhalen;
- ✓ Verwondering weten op te wekken;
- ✓ Meedoen aan gemeenschappelijke activiteiten;
- ✓ Verantwoordelijkheid geven aan kinderen;
- ✓ Initiatieven van de kinderen stimuleren;
- ✓ Met kinderen samen mogelijkheden bekijken, keuzes af wegen en beslissingen nemen.

Bij het begrip criteria moeten we twee elementen benadrukken. Ook de context waarin leerlingen hun capaciteiten moeten demonstreren, kan eventueel in de criteria worden gespecificeerd. Zo kan bijvoorbeeld melding worden gemaakt van de grootte van de groep kinderen, of de leerling de eindverantwoordelijkheid draagt dan wel een kleuterjuf assisteert en of de leerling een hele dag aan deze criteria dient te voldoen. Daarnaast is het onderscheid tussen expliciete en impliciete criteria van belang (Mabry, 1999). Expliciete criteria worden neergeschreven alvorens de kwaliteit van een leerlingprestatie te beoordelen. Van impliciete criteria is sprake wanneer leerkrachten de leerlingprestaties beoordelen op basis van hun observaties en hun professi-

onele ervaring en expertise zonder te verwijzen naar formele richtlijnen. De criteria waarover wij in dit handboek spreken, zijn steeds expliciet.

3.1.2 Wat is een standaard?

Het is belangrijk het begrip criteria te onderscheiden van het begrip standaard. Een standaard wordt geconcretiseerd aan de hand van criteria. Een standaard is bijvoorbeeld dat we van leerlingen verwachten dat ze bij het verlaten van het secundair onderwijs 'goed kunnen schrijven'. Om aan te geven wat we verstaan onder 'goed schrijven' doen we een beroep op criteria. Hierbij denken we bijvoorbeeld aan duidelijkheid, organisatie, levendigheid, gevatheid, uitwerking en innemendheid (Baker et al., 1994). Een standaard is echter meer dan een opsomming van criteria. Aan de hand van voorbeelden, modellen en specificaties wordt in een standaard ook aangegeven hoe duidelijk, levendig en gevat een leerling moet kunnen schrijven (Wiggings, 1998).

Het verband tussen criteria en standaarden kan het best worden toegelicht aan de hand van een voorbeeld zoals hoogspringen. De criteria stellen dat een geldige sprong betekent dat de atleet over de lat springt zonder dat die valt. Dat wil zeggen dat indien de lat beweegt maar blijft liggen toch nog wordt tegemoet gekomen aan de criteria. Het criterium geeft aan wat de hoogspringer moet doen. De vraag blijft echter hoe hoog een geldige sprong moet zijn om als een voldoende of uitmuntende prestatie te gelden. Dat is de vraag naar de standaard. Een standaard geeft aan hoe goed een prestatie moet zijn (Mabry, 1999).

Criteria zijn algemeen, ze gelden over verschillende taken, leerlingen en contexten. Zo moet een schrijfpdracht steeds goed georganiseerd, duidelijk en inzichtelijk zijn. De vraag naar de standaarden is dan: hoe goed is goed genoeg op deze criteria? Uiteraard is het antwoord op deze vraag wel contextafhankelijk. Een essay van een leerling uit de eerste graad van het secundair onderwijs kan volgens de standaarden uitmuntend zijn terwijl hetzelfde essay als product van een laatstejaars duidelijk als ondermaats zal worden beoordeeld. Ondertussen blijven dezelfde criteria echter van tel. Een criterium kan dus uitmonden in verschillende standaarden. Of in het geval van de criteria in figuur 3.5: naargelang het criterium 'initiatieven van de kinderen stimuleren' wordt toegepast in het eerste dan wel in het tweede leerjaar van de derde graad kan het resulteren in andere standaarden. De volgende alinea's concretiseren de begrippen criteria en standaarden.

Een standaard geeft aan welke criteria van belang zijn bij het beoordelen van een leerlingprestatie. In het voorbeeld omtrent 'goed schrijven' zijn de criteria onder andere de 'uitgewerktheid' en de 'organisatie' van het verhaal. Deze criteria moeten verder worden uitgewerkt zodat het mogelijk wordt een geobserveerd gedrag in een prestatieniveau onder te brengen. Het komt er in feite op neer om voor elk prestatieniveau dat je in je beoordeling wil onderscheiden te omschrijven door welke specifieke leerlinggedragingen het gekenmerkt wordt (positieve en negatieve). Pas dan geeft een standaard aan hoe goed een prestatie moet zijn om een bepaalde beoordeling te verdienen. Stel dat ik een schaal met 6 prestatieniveaus onderscheid -gaande van uitzonderlijk

schrijftalent over competente schrijver tot een onvoldoende- dan zal de standaard voor de lovenswaardige schijfprestatie er bijvoorbeeld als volgt uitzien:

Er is een duidelijke inhoudselementen en een logische opeenvolging van inhouden en gebeurtenissen. Er wordt af en toe van de essentie afgeweken, maar dat leidt niet tot afleiding bij de lezer. De overgangen zijn vlot en logisch. De verschillende elementen van het verhaal zijn goed uitgewerkt en er worden verschillende zinsconstructies gehanteerd. De woordenschat is gepast. De leerling blijft niet hangen in algemene stellingen, maar af en toe is de specificiteit wel zoek.

Een dergelijke standaard kan voor elk prestatieniveau worden opgesteld. Naargelang het prestatieniveau zal de invulling van de criteria echter telkens verschillen. Een weergave van de verschillende prestatieniveaus met de bijbehorende standaarden resulteert in een zogenaamde rubriek. Het gebruik van rubrieken lichten we uitvoerig toe in het verdere verloop van deze derde fase in het evaluatieproces (zie paragraaf 3.3).

3.1.3 Wat zijn verwachtingen?

Standaarden zijn op hun beurt niet altijd hetzelfde als verwachtingen. Naargelang de nadruk ligt op wat je verwacht dat leerlingen zouden moeten bereiken, dan wel op wat je verwacht dat ze effectief zullen bereiken, krijgt de term 'verwachting' immers een verschillende invulling. Doorgaans verstaan we onder een verwachting wat de leerlingen op het einde van een leseenheid moeten bereiken. De verwachting komt dan in principe neer op het minimaal tegemoet komen aan de vooropgestelde leerdoelen (zoals de eindtermen). In dat geval is de verwachting gelijk aan de standaard voor een voldoende prestatie. Dat is echter niet altijd het geval. Een evaluatie wil doorgaans een beeld geven van het actuele prestatieniveau van leerlingen. Dit actuele niveau hoeft niet samen te vallen met de vooropgestelde verwachte standaard. De verwachting welk niveau bereikt zal worden, is leerling- of groepsgebonden en kan zowel hoger als lager liggen dan de standaard. De verwachting ten aanzien van de prestaties van een individuele leerling of van een groep leerlingen hoeft dus niet overeen te stemmen met de vooropgestelde standaard. Bijvoorbeeld: ook al is de standaard voor een voldoende prestatie dat leerlingen 26 van de 40 opgaven correct oplossen, toch kan ik van een leerling verwachten dat deze minder dan de helft van de opgaven tot een goed einde zal brengen.

3.2 Het belang van criteria en standaarden

Alternatieve evaluatievormen zijn vaak gebaseerd op observatie en menselijke beoordelingen. Beide zijn evaluatieaspecten waarin subjectiviteit een belangrijke rol kan spelen. Dit houdt een aantal potentiële gevaren in. Wat gebeurt er bijvoorbeeld wanneer de criteria niet consistent door de verschillende beoordelaars worden gebruikt. Of nog: wat gebeurt er indien de beoordelaars hun criteria niet voldoende (helder) verwoorden? Dergelijke vragen maken duidelijk dat een goed gebruik van criteria en standaarden noodzakelijk is. Waarom dat zo is, lichten we toe

in paragraaf 3.2.1. Daarna rapporteren we in paragraaf 3.2.2 een aantal aandachtspunten bij het gebruik van criteria.

3.2.1 De nood aan criteria en standaarden

De noodzaak van evaluatiecriteria en standaarden blijkt zowel uit de positieve gevolgen van de beschikbaarheid ervan als uit de consistentie waartoe ze bijdragen.

1. Beschikbaarheid van criteria en standaarden

Een belangrijke eis is steeds dat de criteria duidelijk geformuleerd en openbaar ter beschikking zijn. De beschikbaarheid van criteria is immers een waardevol instrument in het leerproces van leerlingen. In het pleidooi voor de beschikbaarheid van evaluatiecriteria en -standaarden wordt een onderscheid gemaakt tussen de beschikbaarheid van criteria voor leerlingen en de beschikbaarheid voor ouders.

a. De beschikbaarheid voor leerlingen

Leerkrachten maken voortdurend beoordelingen over leerlingen. Deze beoordelingen zijn - zoals reeds aangehaald - gebaseerd op al dan niet geëxpliciteerde criteria. Als leerkracht heb je dus eigenlijk twee mogelijkheden: je kan de criteria kristalhelder aan de leerlingen duidelijk maken of je kan de leerlingen in het ongewisse laten (Arter, 1996b).

Even stilstaan... 3.5: De beschikbaarheid van evaluatiecriteria voor leerlingen

Twee vragen om te reflecteren over het gebruik van criteria in uw klaspraktijk.

1. Weten uw leerlingen vóór de evaluatie wat er van hen verwacht wordt en welke criteria gehanteerd zullen worden om hun prestaties te beoordelen?
2. Men kan zich de vraag stellen waarom criteria momenteel meestal niet geëxpliciteerd worden: omdat we het moeilijk vinden de criteria te expliciteren, omdat er te weinig tijd is of omdat het belang ervan niet voldoende bekend is? Wat is uw mening?

Vast staat dat bij alternatieve evaluatievormen de criteria wel expliciet zijn en waar mogelijk worden besproken met de leerlingen. Herman et al. (1992) stellen dat het gezamenlijk discussiëren over criteria de leerlingen helpt de standaarden en regels te internaliseren die ze nodig hebben om zelfstandig lerenden te worden. Transparante criteria moeten gaandeweg door de leerlingen aangezien worden als een wezenlijk en natuurlijk onderdeel van het leerproces. Criteria die aangeven wat van een goede prestatie verwacht wordt, zijn richtinggevend voor de leerlingen. Daartegenover staat dat het niet weten wat belangrijk is, het zelfvertrouwen van leerlingen ondermijnt (Arter, 1996b). Het bediscussiëren van criteria helpt leerlingen eveneens de perspectieven van hun leerkrachten, hun medeleerlingen en eventueel van experts in het vakgebied te ontdekken. Ook wanneer van leerlingen verwacht wordt

dat ze zichzelf en/of medeleerlingen evalueren, spreekt het voor zich dat de evaluatiecriteria beschikbaar en duidelijk moeten zijn. Kortom: er zijn verschillende redenen om leerkrachten aan te zetten hun evaluatiecriteria naar de leerlingen toe te communiceren. Criteria zijn niet alleen van belang bij het evalueren van de output, ook tijdens het leer- en instructieproces spelen ze een wezenlijke rol.

b. Beschikbaarheid voor ouders

Ook het belang van de beschikbaarheid van criteria voor ouders moet benadrukt worden (Herman et al, 1996b). Dit belang komt neer op het feit dat criteria ook voor ouders duidelijkheid kunnen scheppen over wat leerkrachten en de school bij de leerlingen trachten te bereiken. Wanneer ouders, vooraleer leerlingen geëvalueerd worden, weten wat er van hun kinderen verwacht wordt, kunnen zij het leerproces beter ondersteunen.

Zorgvuldig ontwikkelde criteria helpen dus zowel leerlingen als ouders om mee verantwoordelijkheid te dragen voor het leren. Bij het activeren van leerlingen in hun eigen leer- en evaluatieproces spelen criteria en standaarden een cruciale rol (zie deel 4). Bovendien kan vermeden worden dat onduidelijkheid over de leerinhouden en prioriteiten leidt tot demotivatie bij de leerlingen en tot mogelijke conflicten tussen leerlingen en leerkrachten.

2. Consistentie

Onder consistentie verstaan we dat de waardering van de prestaties van leerlingen voldoende stabiel is over verschillende evaluatiemomenten en verschillende beoordelaars. Eenzelfde prestatie moet op verschillende meetmomenten tot een gelijkaardige score leiden (gegeven dat er ondertussen geen instructie is geweest of groei is opgetreden). Hiertoe is het noodzakelijk criteria en standaarden aan te wenden. Het gebruik van goede criteria draagt er tevens toe bij dat een waardering eveneens stabiel is over verschillende beoordelaars. In dat geval spreken we van interbeoordelaarsconsistentie. Om dat te garanderen moeten de criteria een mechanisme vormen dat overeenstemming tussen beoordelaars creëert en dat duidelijk de essentiële aspecten van de leerlingprestatie beschrijft (Herman et al., 1992; Arter, 1996b). Dat leunt sterk aan bij wat we bespraken bij de betrouwbaarheid als kenmerk van een goede evaluatie. Criteria en standaarden zijn dan ook noodzakelijk om een betrouwbare evaluatie op poten te zetten.

3.2.2 Enkele aandachtspunten bij het gebruik van criteria

In de evaluatieliteratuur is er een grote eensgezindheid over het feit dat expliciete evaluatiecriteria noodzakelijk zijn om tot een goede en betrouwbare evaluatie te kunnen komen. Ook in dit boek wordt de noodzaak van dergelijke criteria en standaarden beklemtoond. Dit neemt niet weg dat het toch belangrijk is het gebruik van expliciete criteria kritisch te bekijken (Mabry, 1999). Om die reden rapporteren we een drietal potentiële problemen bij het gebruik van expliciete criteria waarvan men zich als gebruiker goed bewust moet zijn. Op die manier kunnen eventuele negatieve aspecten tot een minimum herleid worden.

- ✓ Het gebruik van expliciete criteria kan impliceren dat de aandacht van de observator/evaluator te zeer wordt afgeleid van de eigenlijke leerlingprestatie naar de vooropgestelde criteria en standaarden. Op die manier zou de aandacht voor particuliere kenmerken van de leerlingprestatie kunnen afnemen. Tot op zekere hoogte is dit geen probleem, vermits wat we belangrijk vinden in principe vervat zit in de criteria. Dit mag echter niet betekenen dat we enkel de aspecten van een leerlingprestatie die rechtstreeks gerelateerd zijn aan de criteria en standaarden, in acht zouden nemen.
- ✓ Een hieraan gerelateerd probleem is het mogelijk tekort aan flexibiliteit wanneer de criteria te strikt gebruikt worden. Het unieke en verscheiden werk van leerlingen hoeft immers niet per se te passen in de criteria die vooraf werden vooropgesteld. Op die manier zouden goede prestaties toch ondergewaardeerd kunnen worden.
- ✓ Expliciete criteria zouden er verder toe kunnen leiden dat verwacht wordt dat alle leerlingprestaties conform de criteria dienen te zijn. Dit zou kunnen resulteren in een standaardisatie en convergentie in het denken en in de producten van leerlingen waarbij er geen ruimte meer wordt geboden aan creativiteit en zelfexpressie. Leerlingen zouden hun prestaties conformeren aan de vooropgestelde criteria (Mabry, 1999). In dit geval zou er sprake zijn van een variatie op het gekende 'teaching to the test', namelijk 'teaching to the criteria'.

Bij het gebruik van alternatieve evaluatievormen mag men de ogen niet sluiten voor de problemen die kunnen voortvloeien uit het gebruik van criteria. We spreken van potentiële problemen omdat we van mening zijn dat een zorgvuldige ontwikkeling van criteria een stevige garantie kan bieden voor het vermijden ervan. Criteria kunnen samen met leerlingen ontwikkeld worden zodat alle relevante aspecten opgenomen worden. Ze kunnen bovendien zo ontworpen worden dat de creativiteit en flexibiliteit niet in het gedrang komen. Evaluatie van de criteria zelf maakt dat problemen gedetecteerd worden en in een volgende versie kunnen worden weggewerkt. En ten slotte: is er een probleem met het zogenaamde 'teaching to the criteria' als men over een goed doordachte en aangepaste set van criteria beschikt?

Uit deze opmerkingen onthouden we voornamelijk dat het gebruik van criteria en standaarden enkel een meerwaarde zal hebben indien het ontwikkelingsproces en het gebruik ervan met de nodige nauwgezetheid en ernst plaatsvindt.

3.3 Hoe kan je criteria en rubrieken ontwikkelen?

In de resterende paragrafen van de derde stap van het evaluatieproces rapporteren we nog twee zaken. In deze paragraaf beschrijven we eerst hoe je tewerk zou kunnen gaan bij het ontwikkelen van criteria en rubrieken. Daarna staan we stil bij een aantal kwaliteitseisen die ten aanzien van criteria en rubrieken kunnen worden vooropgesteld (zie paragraaf 3.4).

Ook in het ontwikkelen van criteria kunnen een aantal stappen worden onderscheiden. Elk van deze stappen wordt in de volgende beschrijving uiteengezet.

Figuur 3.6: Stappen bij het ontwikkelen van evaluatiecriteria

1. zet de verschillende leeruitkomsten (dimensies) op een rij;
2. verzamel voorbeelden van deze leeruitkomsten;
3. groepeer de voorbeelden in verschillende prestatieniveaus;
4. beschrijf in rubrieken de karakteristieken voor elk prestatieniveau;
5. ontwikkel een schaal om de mate waarin een bepaald niveau bereikt werd te beoordelen.

3.3.1 Zet de verschillende leeruitkomsten (dimensies) op een rij

Alvorens criteria te kunnen ontwikkelen, moeten de leerlingenprestaties die geëvalueerd zullen worden, duidelijk afgeijnd zijn. Het op een rij zetten van de verschillende leeruitkomsten is dan ook de eerste stap.

De prestaties van leerlingen bij levensechte vaardigheidstaken omvatten meestal meerdere vaardigheden. Kenmerkend voor deze taken is immers net dat verschillende vaardigheden gecombineerd dienen te worden. Om tot valide en informatieve evaluatiegegevens te komen, is het echter van belang dat de beoordeling van al deze vaardigheden afzonderlijk gebeurt. Daarom is het zinvol in het te evalueren domein verschillende leeruitkomsten (zogenaamde dimensies) te onderscheiden. Een leeruitkomst van een evaluatietaak kunnen we omschrijven als nagestreefde kennis of als een te ontwikkelen (deel)vaardigheid. Uiteraard is het eveneens mogelijk dat een evaluatietaak zich enkel toelegt op één bepaalde vaardigheid. In dat geval is deze eerste stap in principe overbodig.

Om het begrip leeruitkomst te illustreren, geven we twee korte voorbeelden.

Voorbeeld 1: Vaak wordt de schrijfvaardigheid van leerlingen in één omvattend cijfer beoordeeld. Men gaat er in dat geval van uit dat het in de vaardigheid schrijven niet nodig of mogelijk is verschillende dimensies te onderscheiden. Deze stelling is echter niet correct. Met het oog op het ontwikkelen van betere evaluatie-instrumenten onderscheiden Spandel & Culham (1996) binnen het leerdoel 'schrijven' volgende zes dimensies:

- ✓ ideeën en inhoud,
- ✓ organisatie,
- ✓ stijl,
- ✓ woordkeuze,
- ✓ vlotheid,
- ✓ taalconventies.

Voorbeeld 2: Laten we ervan uitgaan dat we met een evaluatie de vaardigheden van een leerling willen beoordelen om in groep te werken. Ook binnen dit ruimere evaluatiedomein kunnen we een zestal dimensies onderscheiden:

- ✓ aanpassingsvermogen: problemen herkennen en gepast beantwoorden;
- ✓ coördinatie: de groepsactiviteiten organiseren om tijdig de taak te volbrengen;
- ✓ beslissingen nemen: de beschikbare informatie gebruiken om beslissingen te nemen;
- ✓ interpersoonlijk: coöperatief interageren met de andere groepsleden;
- ✓ leiderschap: sturing en richting geven aan de groep;
- ✓ communicatie: duidelijk en accuraat informatie uitwisselen tussen groepsleden.

In de dimensies die men kiest, moeten de belangrijkste kwaliteiten die men bij leerlingen wenst te bereiken, weerspiegeld worden. Hierbij moet men zich trachten te focussen op doelstellingen die effectief te onderwijzen en te observeren zijn.

Eenmaal de verschillende dimensies (leeruitkomsten) gekend zijn, kan het ontwikkelingsproces van criteria worden verdergezet.

3.3.2 Verzamel voorbeelden van de betreffende leeruitkomsten

Het is aangewezen voor elk van de onderscheiden leeruitkomsten (dimensies) van de taak voorbeelden te verzamelen. Afhankelijk van de vooropgestelde onderwijsdoelen kan een dergelijke verzameling bijvoorbeeld bestaan uit verschillende probleemoplossingsmethoden, verschillende opstellen, verschillende opgenomen gesprekken van leerlingen met rusthuisbewoners en verschillende voorbeelden waarin leerlingen de nodige veiligheidsvoorschriften toepassen. Indien men als leerkracht (nog) beschikt over de prestaties van leerlingen bij vroegere evaluaties (bijvoorbeeld de opstellen van de voorbije schooljaren) kunnen deze handig aangewend worden. In het geval er geen voorbeelden beschikbaar zijn, moeten die nog verzameld worden. Dit kan door een aantal leerlingen de betreffende evaluatietaak te laten uitvoeren of door zelf op basis van eigen ervaringen en observaties een aantal voorstellen te beschrijven.

Bij het verzamelen van voorbeelden is het belangrijk een verscheidenheid aan prestatieniveaus na te streven, zodat zowel schitterende, voldoende als onvoldoende prestaties van de verzameling deel uitmaken. Wanneer men beschikt over voldoende voorbeelden van antwoorden op de betreffende evaluatietaak bestaat de volgende stap uit het aanbrengen van een ordening.

3.3.3 Groepeer de voorbeelden en voorzie voor elke leeruitkomst verschillende prestatieniveaus

In deze stap van het ontwikkelingsproces van criteria komt het erop aan de beschikbare voorbeelden in een aantal groepen (kwaliteitsbeoordelingen) te verdelen. Hiervoor baseert men zich op beoordelingsschalen. Schalen bestaan in verschillende vormen en structuren (zie stap e). Een mogelijkheid is een onderverdeling in vijf prestatieniveaus. De beschikbare voorbeelden zouden in dit geval met andere woorden in vijf kwaliteitsgroepen verdeeld kunnen worden:

-de minimale prestatie	-onvoldoende
-de rudimentaire prestatie	-voldoende met leemtes
-de lovenswaardige prestatie	-voldoende
-de uitstekende prestatie	-goed
-de buitengewone prestatie	-schitterend

Voor elk van de te evalueren dimensies kan men deze prestatieniveaus onderscheiden. Het aantal niveaus dat men onderscheidt, kan uiteraard variëren en moet zorgvuldig overwogen worden. Voornamelijk de beslissingen die men op basis van de evaluatiegegevens wil nemen en de gedetailleerdheid waarmee men de informatie wil terugkoppelen, zijn van belang bij het kiezen van het aantal niveaus. Het aantal moet groot genoeg zijn met het oog op de effectiviteit en klein genoeg met het oog op de betrouwbaarheid en beheersbaarheid.

Bij het sorteren van de beschikbare voorbeelden is het van belang na te gaan waarom een bepaald voorbeeld in een bepaalde kwaliteitsgroep ondergebracht wordt. Je moet je met andere woorden afvragen welke kenmerken van een voorbeeld maken dat het bijvoorbeeld in de groep 'goede prestatie' thuishoort. En verder: waarin verschilt de goede prestatie van de onvoldoende prestatie? Door deze evaluatieve commentaren neer te schrijven, ontdek je potentiële evaluatiecriteria. De aard, de inhoud en de hoeveelheid van de relevante kenmerken, en dus van de toekomstige criteria, kunnen sterk variëren. Bij wijze van illustratie sommen we een aantal commentaren op die relevant kunnen zijn om het onderscheid te maken tussen verschillende kwaliteitsniveaus van een door leerlingen uitgevoerde zelfreflectie: omvattendheid, gedetailleerdheid, motivatie, vooruitblikken, sterkte- en zwakteanalyse, kwaliteit van de verklaringen en voorbeelden, vergelijking over de tijd, procesgerichtheid, de mate waarin doelen voor de toekomst vooropzet worden, leesbaarheid, gestructureerdheid, eerlijkheid en oprechtheid, etc.

Even recapituleren. Een leeruitkomst omschreven we als een vorm van kennis of een vaardigheid die je van leerlingen verwacht en waaraan je dus aandacht zal besteden tijdens de evaluatie. Eén van de leeruitkomsten binnen de vaardigheid 'goed schrijven' is het tegemoet komen aan taalconventies. Als je wat deze leeruitkomst betreft een vijftiental opstellen onderverdeelt in vier prestatieniveaus, kan je door de verschillen tussen de prestatieniveaus te motiveren criteria op het spoor komen. Criteria zijn de voorwaarden waaraan een (leerling)prestatie dient te voldoen om als succesvol te worden bestempeld. In het geval van het tegemoetkomen aan schrijfconventies zullen de verschillen tussen prestatieniveaus bijvoorbeeld terug te brengen zijn tot volgende criteria: het naleven van de grammatica- en spellingsregels, het correct gebruik van leestekens. Op een zelfde manier kunnen criteria worden ontwikkeld voor de andere dimensies van 'goed kunnen schrijven'.

Figuur 3.7: Mogelijke criteria voor de dimensies van 'goed schrijven'

- ✓ ideeën en inhoud: duidelijkheid, gevatheid, relevantie, doelgerichtheid;
- ✓ organisatie: logische en vloeiende structuur, uitnodigende introductie, goed besluit;
- ✓ stijl: interactie met de lezer, eerlijk, overtuigend, vragen anticiperen, smaakmakende toon;

- ✓ woordkeuze: specifiek, natuurlijk, accuraat, levendig en niet overladen;
- ✓ vlotheid: duidelijke betekenisvolle zinnen, adequaat gebruik van voegwoorden en adjectieven;
- ✓ schrijfconventies: correct gebruik van grammatica en spelling, goed gebruik van paragrafen en leestekens.

Hiermee wordt duidelijk dat het onderscheid tussen leeruitkomsten en criteria een kwestie is van invalshoek. Vast staat alvast dat criteria aangeven waaraan een leeruitkomst moet voldoen om succesvol te zijn. Dit houdt echter in dat naargelang de invalshoek die je gebruikt leeruitkomsten criteria kunnen zijn en criteria opnieuw leeruitkomsten. Zo kan je de leeruitkomsten 'ideeën en inhoud', 'organisatie', 'stijl' en 'woordkeuze' aanzien als criteria voor de leeruitkomst 'goed schrijven' en kan 'correct gebruik van grammatica' aangezien worden als een leeruitkomst waarvoor volgens het voorgestelde stappenplan opnieuw criteria ontwikkeld kunnen worden. Het hangt er dus vanaf wat de focus is van je evaluatie en hoe gedetailleerd je gegevens wil verzamelen. Vandaar het belang om in de eerste stap van het ontwikkelen van criteria duidelijk te omschrijven wat de leeruitkomsten zijn waarover je uitspraken wil doen en waarvoor je dus criteria wil ontwikkelen.

Bij het groeperen van de verzamelde voorbeelden en het voorzien van prestatieniveaus en criteria voor elke leeruitkomst zijn nog twee opmerkingen relevant. Beide hebben onrechtstreeks betrekking op de band tussen leeruitkomsten en evaluatiecriteria.

- ✓ De verbondenheid tussen leeruitkomsten en criteria maakt dat er enige variatie kan ontstaan in het ontwikkelingsproces van criteria. Door expliciet aandacht te besteden aan de kwaliteitskenmerken van voorbeeldprestaties kan duidelijk worden dat binnen de leeruitkomst waarvoor voorbeelden verzameld worden nog verschillende leeruitkomsten (dimensies) onderscheiden kunnen worden. Het sorteren van voorbeelden van complexe, samengestelde vaardigheden (zoals bijvoorbeeld de vaardigheid in groep te kunnen werken) kan dus ook als strategie aangewend worden om de verschillende relevante dimensies en leeruitkomsten aan het licht te brengen. Op die manier komt men via criteria tot leeruitkomsten. Bij het groeperen van bijvoorbeeld schrijfproduct in verschillende kwaliteitsgroepen komen als onderscheidende argumenten wellicht criteria zoals structuur, woordgebruik, originaliteit, correct taalgebruik en gevatheid naar voren. Al deze elementen zouden bij het beoordelen van een opstel in rekening moeten worden gebracht en zijn dus belangrijke leeruitkomsten. Kortom: door de kwaliteitsverschillen tussen de antwoorden te verantwoorden kan men belangrijke leeruitkomsten (kennis en vaardigheden) op het spoor komen.
- ✓ Indien er bij het verantwoorden van verschillen in prestatieniveaus teveel criteria onderscheiden kunnen worden, komt het er op aan deze verschillende kwaliteitskenmerken te clusteren tot overkoepelende leeruitkomsten (Arter, 1996b). Het aantal criteria moet immers behapbaar blijven en dat betekent dat men indien nodig de leeruitkomsten op een hoger niveau als criteria moet hanteren. In het opstelvoorbeeld kunnen binnen die gedachtegang criteria zoals correcte grammatica, correcte spelling en goed gebruik van paragrafen en leestekens worden samengebracht in de

omvattende leeruitkomst 'tegemoetkomen aan taalconventies'. Een ander voorbeeld vindt u in volgende denkoefening.

Even stilstaan... 3.6: Criteria clusteren tot leeruitkomsten

De criteria die we gebruiken om het verschil tussen een goede en een minder goede zelfreflectie te verantwoorden, zijn onder andere: omvattendheid, gedetailleerdheid, motivatie, vooruitblikken, sterkte- en zwakteanalyse, goede verklaringen, vergelijking over de tijd, procesgericht, voorbeelden geven, doelen vooropzetten voor de toekomst, leesbaar, goed gestructureerd, eerlijk en oprecht.

Sommige van deze criteria kunnen geclusterd worden tot ruimere leeruitkomsten. Zo kunnen de criteria 'vooruitblikken' en 'doelen vooropstellen voor de toekomst' samengebracht worden in de dimensie doelen vooropstellen. Binnen de aangegeven kenmerken van zelfreflectie kunnen echter nog andere leeruitkomsten onderscheiden worden. Kan u nagaan welke resterende criteria geclusterd kunnen worden tot de dimensies 'analyse van de eigen vaardigheden', 'oprechtheid' en 'kunnen presenteren'? Ziet u nog andere dimensies?

Wat ook de gehanteerde volgorde is in het ontwikkelen van criteria en leeruitkomsten, vast staat dat men aan het einde van deze fase moet beschikken over een aantal criteria die het mogelijk maken een onderscheid te maken in verschillende prestatieniveaus van leerlingen (en dit voor elk van de onderscheiden leeruitkomsten en dimensies). Vervolgens worden in een volgende stap aan de hand van deze criteria en de keuze van een aantal prestatieniveaus de karakteristieken van de verschillende niveaus besproken. Dit uit zich in een zogenaamde rubriek.

3.3.4 Beschrijf in rubrieken de karakteristieken voor elk prestatieniveau

Wanneer voor de relevante leeruitkomsten verschillende prestatieniveaus zijn vooropgesteld, bestaat de volgende stap erin voor elk prestatieniveau de relevante karakteristieken te omschrijven. Het komt er op neer voor elk prestatieniveau duidelijk te omschrijven door welke leerlinggedragingen (positieve en negatieve) het gekenmerkt wordt. Dit betekent dat je aan de hand van de in de vorige stap ontwikkelde criteria aangeeft waaraan een geobserveerd gedrag of een product moeten voldoen om ze in een bepaald prestatieniveau onder te brengen.

Alvorens het begrip 'rubriek' te definiëren en verder te omschrijven, geven we in de volgende figuur een illustratie.

Figuur 3.8: Een rubriek ter illustratie

Het voorbeeld richt zich op de actieve en passieve mondelinge Engelse taalbeheersing. In deze vaardigheid kunnen verschillende criteria (of leeruitkomsten) onderscheiden worden: taalbegrip, communicatie, vloeiendheid, structuur en woordenschat. De rubriek die wij ter illustratie voorstellen, richt zich specifiek op het taalbegrip, op het evalueren van de capaciteiten om gesproken Engels te ver-

staan. Een zelfde rubriek zou men echter kunnen opstellen voor de andere criteria. In de rubriek voor taalbegrip worden zes prestatieniveaus onderscheiden.

- ✓ buitengewoon: verstaat zonder enige moeite academische uiteenzettingen;
- ✓ zeer goed: verstaat de meeste vormen van gesproken Engels;
- ✓ goed: verstaat gesproken Engels mits herhalingen, herformuleringen en verduidelijkingen;
- ✓ voldoende: verstaat eenvoudige gesproken zinnen in een conversatie mits herhalingen;
- ✓ pover: verstaat woorden en zinnen mits herhalingen;
- ✓ slecht: verstaat weinig of geen Engels.

Het begrip rubriek omschrijven we aan de hand van de volgende twee definities: (1) Een rubriek is een geschreven geheel van richtlijnen dat het mogelijk maakt prestaties en producten van verschillende kwaliteiten te onderscheiden. Een rubriek omvat beschrijvingen die aangeven waarnaar men moet kijken bij elk kwaliteitsniveau. Ze maken duidelijk wat een bepaald prestatieniveau kenmerkt. Op die manier blijkt welke aspecten van de prestatie of het product het meest belangrijk zijn (Wiggins, 1996). (2) Een rubriek is een evaluatieinstrument dat ontwikkeld wordt in de vorm van een kwaliteitscontinuüm gaande van excellent tot zeer slecht. Elk niveau van de rubriek omvat een narratieve beschrijving van hoe een prestatie of een product dat het betreffende niveau behaalt er dient uit te zien (ASCD, 1996). Een rubriek is dus een overzicht van welke leerlingprestaties kenmerkend zijn voor de verschillende prestatie- of beoordelingsniveaus. Het is aan de beoordelaar om te beslissen aan welke criteria de geobserveerde leerlingprestatie al dan niet voldoet.

In principe zou je een rubriek kunnen opstellen voor elk criterium dat in vorige stap onderscheiden werd. Voor elk criterium kan immers hetzelfde onderscheid naar prestatieniveaus gemaakt worden. Dit betekent bijvoorbeeld dat je in het voorbeeld van goed schrijven telkens een rubriek ontwerpt voor duidelijkheid, gevatheid, relevantie, doelgerichtheid, structuur, interactie met de lezer, eerlijkheid, overtuigendheid, toon, levendigheid enzoverder. Voor het beoordelen van een essay zouden dan meer dan 20 rubrieken ontworpen en gebruikt moeten worden. Hoewel op die manier zeer goede feedback gegenereerd zou kunnen worden, is dit niet werkbaar. Het komt er dan ook op aan verschillende leeruitkomsten en criteria samen te voegen in één rubriek. In het geval van goed schrijven zou dit bijvoorbeeld kunnen leiden tot twee rubrieken, de kwaliteit van de ideeën en argumenten en de kwaliteit van het schrijven op zich.

De voordelen van een rubriek zijn gelijkaardig aan de voordelen verbonden aan het gebruik van evaluatiecriteria. Het belangrijkste doel van het gebruik van rubrieken is dan ook vooraf duidelijk te maken welke leeruitkomsten nagestreefd worden (Conley, 1996). Dit komt deels overeen met wat we benadrukten met de beschikbaarheid en de consistentie die nagestreefd wordt door het gebruik van evaluatiecriteria. Ook de mogelijkheden die rubrieken bieden bij formatieve evaluatie worden in de evaluatieliteratuur benadrukt.

Rubrieken bestaan in verschillende soorten. Ze kunnen zowel samengesteld worden met het oog op het scoren van antwoorden op één enkele taak als met het oog op het beoordelen van de prestaties van leerlingen over verschillende taken. Daarnaast wordt nog een onderscheid gemaakt tussen analytische en holistische rubrieken. Het begrip 'rubriek' en het onderscheid tus-

sen de verschillende types kunnen best geduid worden aan de hand van een aantal concrete voorbeelden. We maken hierbij zowel een onderscheid tussentaakgebonden en algemene rubrieken als tussen analytische en holistische rubrieken (Arter, 2000; Moskal et al., 2000; Wiggins, 1998).

1. Taakgebonden en algemene rubrieken

De taakgebonden rubriek heeft tot doel de antwoorden van leerlingen op één specifieke taak te beoordelen. Het eigene aan deze rubrieken is dat ze taakspecifiek zijn in die zin dat ze niet kunnen gebruikt worden om de antwoorden van leerlingen op een andere evaluatie-taak te beoordelen. Uit de aard van de taakgebonden rubriek blijkt dat deze soort rubriek vaak niet beantwoordt aan de kenmerken van rubrieken en evaluatiecriteria waarvan we positieve effecten verwachten. We hebben deze vorm van rubrieken toch opgenomen om aan te tonen dat voor elke taak duidelijke scoringsrichtlijnen opgesteld kunnen worden die een betrouwbare beoordeling mogelijk maken, waarbij subjectiviteit vermeden wordt. Bij het gebruik van een taakgebonden rubriek moet je daartoe als beoordelaar nauwgezet de vooropgestelde scoringsrichtlijnen volgen. De samenstellers van de rubriek hebben reeds beslist waarop de beoordelaar moet letten en wat hij moet scoren als een goede of een ontoereikende prestatie (Arter, 1996b). Dit blijkt uit het voorbeeld van een taakgebonden rubriek in figuur 3.9.

Figuur 3.9: Voorbeeld van een taakgebonden rubriek

De evaluatietaak: Aan de hand van de tabel van Mendeljev wordt de leerlingen gevraagd acht chemische elementen te categoriseren en te rangordenen (de elementen zijn Li, Cs, Fe, Ti, Al, P, S & Cl). De opdrachten zijn:

1. Sorteer de elementen aan de hand van volgende chemische families: alkalimetalen, andere metalen en niet-metalen.
2. Bereken voor elk element de atoommassa.
3. Rangschik de elementen van klein naar groot naargelang hun atoommassa. Zie je een verband tussen de atoommassa en de plaats van het element in de Tabel van Mendeljev. Schrijf dit verband in één zin neer.

Mogelijke rubriek (bij scoring van de taak op 10 punten)

In het antwoord worden logischerwijs drie verschillende dingen gescoord. (1) Sorteren: 4 punten indien alle elementen correct gesorteerd zijn. Trek hiervan per foutief gesorteerd antwoord een half punt af. (2) Atoommassa berekenen: 4 punten indien alle atoommassa's correct berekend zijn. Trek hiervan per foutief gesorteerd antwoord een half punt af. (3) Verband: 2 punten indien het antwoord verwijst naar het verband tussen het atoomnummer en de atoommassa; 1 punt indien het antwoord melding maakt van linksboven en rechtsonder; in andere gevallen 0 punten.

De beoordeling berust bij taakgebonden rubrieken niet altijd op het gebruik van evaluatiecriteria. Meestal wordt aan de antwoorden een score toegekend op basis van de aan- of afwezigheid van bepaalde woorden, zinnen, berekeningen, schema's en dergelijke. Een dergelijke taakgebonden rubriek mag dan wel consistentie over meetmomenten en beoordelaars garanderen, de informatieve waarde van de rubriek voor de leerling en leerkracht is eerder gering. Wat leert een score van 5 mij bijvoorbeeld over de vraag waaraan ik in het verdere

verloop van het leerproces bijkomende aandacht moet besteden. Volgende rubriektypes sluiten beter aan bij de assessmentcultuur.

Niet-taakgebonden rubrieken (of rubrieken voor algemene beoordelingen) zijn in een tweetal opzichten verschillend van de vorige groep: (1) de beoordelaars moeten zelf nog een kwalitatief oordeel uitspreken over de prestaties van de leerlingen en (2) de criteria kunnen gebruikt worden voor meerdere oefeningen en evaluatietaken. Een belangrijk voordeel van rubrieken voor algemene beoordelingen ten opzichte van taakgebonden rubrieken is dat ze een beter beeld geven van hoe een goede prestatie er uit ziet (Arter, 1996b). Ze geven een ruimere omschrijving van de nagestreefde prestatie (doelstelling) en beperken zich niet tot het aangeven van hoe de prestatie zich bij één concrete taak dient te manifesteren. Op basis van taakgebonden rubrieken kan je niet afleiden wat een goede prestatie is op een andere taak, en in dat opzicht zijn dergelijke rubrieken dus minder leerrijk, minder stimulerend en minder geschikt als basis voor (zelf)reflectie en feedback. Het gebruik van 'algemene' terugkerende criteria heeft dus een sterkere impact op het leren en de instructie. Ze genieten dan ook -indien mogelijk- de voorkeur boven taakspecifieke criteria.

Volgend voorbeeld geeft een beeld van hoe rubrieken voor algemene beoordelingen eruit zien.

Figuur 3.10: Voorbeeld van een algemene niet-taakgebonden rubriek

Volgende rubriek werd samengesteld met het oog op de (zelf)evaluatie van leerlingen secundair onderwijs¹. De rubriek richt zich weliswaar op één bepaalde leerdoel, namelijk het maken van grafieken, maar kan gebruikt worden bij het beoordelen van verschillende evaluatietaken. Er worden 6 prestatieniveaus onderscheiden:

- ✓ *A - Buitengewoon: De grafiek van de leerling is uitmuntend in het duidelijk en gemakkelijk overdragen van accurate informatie. De grafiek is zeer interessant.*
- ✓ *B - Voortreffelijk: De leerling heeft een aan de data aangepast type grafiek gekozen. De variabelen zijn op de correcte assen geplaatst en de labels zijn correct. Er werd een geschikte schaal met zowel een aangepast startpunt als een aangepast interval gebruikt. De grafische voorstelling van de data is accuraat en aanwezige trends (of de afwezigheid daarvan) worden duidelijk aangegeven. De titel is duidelijk en beschrijft de verschillende variabelen. Technieken zoals kleur en patronen maken dat de grafiek onmiddellijk duidelijk is. De grafiek is ordelijk en presenteerbaar.*
- ✓ *C - De grafiek zou de beoordeling 'voortreffelijk' krijgen ware het niet dat één of twee belangrijke elementen niet voortreffelijk zijn.*
- ✓ *D - De grafiek zou de beoordeling 'pover' krijgen ware het niet dat één of twee belangrijke elementen goed zijn.*
- ✓ *E - Pover: De leerling heeft een niet geschikt of incorrect type grafiek gekozen. De variabelen zijn op de verkeerde assen geplaatst. De assen zijn niet of inadequaet benoemd. De gebruikte schaal is niet geschikt en de lijnen of balken gebruiken de beschikbare ruimte niet*

¹ Gebaseerd op Hibbard, 1996.

goed. De titel ontbreekt of is foutief. Technieken zoals kleuren en patronen werden niet gebruikt om de correcte interpretatie van de grafiek te bevorderen. De grafiek oogt slordig.

✓ *F - Zeer slecht: de grafiek is zeer slecht.*

Uit dit voorbeeld blijkt dat algemene rubrieken gebruikt kunnen worden om verschillende evaluatietaken omtrent het opstellen van grafieken te beoordelen.

2. Analytische en holistische rubrieken

Rubrieken kunnen ook onderverdeeld worden naargelang ze analytisch dan wel holistisch zijn. **Analytische rubrieken** delen de prestatie van een leerling als het ware op in zoveel prestaties als er criteria zijn. Het tegemoet komen aan één van de criteria wordt aangezien als een soort deelprestatie die afzonderlijk beoordeeld moet worden. Analytische rubrieken voorzien dan ook feedback over de verschillende leeruitkomsten en criteria. Deze rubrieken mogen dan wel minder efficiënt zijn in hun gebruik dan holistische rubrieken, ze genereren meer gedetailleerde feedback die een gerichtere planning van het verdere leer- en instructieproces mogelijk maakt. Dit blijkt uit het voorbeeld van een analytische rubriek in figuur 3.11.

Figuur 3.11: Een analytische rubriek voor het geven van een mondelinge presentatie

	ORGANISATIE	INHOUD	OVERDRACHT	MEDIA
EX-CEL-LENT	De vooruitgang is excellent. De luisteraar is steeds mee. Vlotte overgangen.	Uitvoerige, volledige behandeling van het onderwerp. Getuigt van onderzoek.	Goed oogcontact, goede intonatie en lichaamstaal. Meeslepend en genietbaar.	Excellent gebruik van overhead, video's, grafieken. Duidelijk en levendig.
GOED	Goede vooruitgang. De luisteraar blijft bij de inhoud.	Goede dekking van het onderwerp met evidentie van onderzoek.	Oogcontact, intonatie, lichaamstaal en gedrag zijn goed maar niet consistent.	Visualiserings zijn duidelijk en geschikt. De media ondersteunen de presentatie.
VOL-DOEN-DE	Soms wat stroef. De luisteraar is soms verward. Inhoud gaat van de hak op de tak.	Het onderzoek blijkt wel, maar het is niet uitgebreid of er ontbreekt diepgang. Enigszins gedocumenteerd.	Weinig aansprekende overdracht. De presentatie wordt afgelezen of lijkt mechanisch. Niet slecht maar ook niet goed.	Er worden visualisaties gebruikt maar deze zijn niet echt waardevol. Ze helpen de presentatie niet echt vooruit.
ON-VOL-DOEN-DE	Zeer moeilijk te volgen. De luisteraar heeft het moeilijk de aandacht erbij te houden. Onduidelijke overgangen.	Het onderzoek lijkt beperkt. Weinig documentatie. Belangrijke bronnen ontbreken.	De gegevensoverdracht is van een slechte kwaliteit. Het levert niet meer op dan een lezing.	De visualisaties zijn van een dergelijk slechte kwaliteit dat ze de presentatie in het geheel niet ondersteunen.

Holistische rubrieken richten zich op het maken van globale beoordelingen van de leerlingprestaties. Vaak wordt de rubriek gebruikt om uitspraken over verschillende evaluatietaken heen te genereren, maar dat hoeft niet noodzakelijk het geval te zijn. De essentie bestaat erin dat men in één beoordeling een uitspraak doet over de leerlingprestatie en dit op basis

van een omvattende indruk van het werk van de leerling. Holistische beoordelingen zijn efficiënt en snel. Daartegenover staat dat niet voorzien kan worden in een eventuele vraag naar gedetailleerde informatie ten aanzien van specifieke criteria. Dat dit in sommige gevallen voor enige problemen kan zorgen, blijkt uit het voorbeeld van een holistische rubriek in figuur 3.12 (uit Wiggins, 1998). Wat zou er bijvoorbeeld moeten gebeuren wanneer een leerling erin slaagt een goed oogcontact te behouden met het publiek terwijl de bestudeerde inhoud en het aangeven van het belang ervan te wensen overlaten?

Figuur 3.12: Een holistische rubriek voor het geven van een mondelinge presentatie

- ✓ 5 - Excellent. De leerling beschrijft duidelijk de bestudeerde inhoud en geeft uitdrukkelijk het belang ervan aan. Er wordt specifieke informatie gegeven die de gemaakte conclusies ondersteunt. De inhoud wordt op een uitnodigende manier gebracht en de zinsbouw is consistent correct. De leerling maakt oogcontact met de aanhouders tijdens gans de uiteenzetting. De presentatie getuigt zeer duidelijk van een goede voorbereiding, organisatie en enthousiasme. Visuele hulpmiddelen dragen bij tot de effectiviteit van de presentatie. Publieksvragen worden duidelijk beantwoord met specifieke en relevante informatie.
- ✓ 4 - zeer goed. De leerling beschrijft de bestudeerde inhoud en geeft het belang ervan aan. Er wordt voldoende informatie gegeven om de conclusies te ondersteunen. Het overbrengen van informatie en de zinsbouw zijn in het algemeen correct. De presentatie getuigt van een goede voorbereiding, organisatie en enthousiasme. Visuele hulpmiddelen worden aangehaald en goed gebruikt. Publieksvragen worden duidelijk beantwoord.
- ✓ 3 - goed. De leerling beschrijft de bestudeerde inhoud en trekt conclusies, maar de aangebrachte informatie is niet zo sterk als in 4 en 5. Het overbrengen van informatie en de zinsbouw zijn in het algemeen correct. De presentatie getuigt van voorbereiding en organisatie. Visuele hulpmiddelen worden aangehaald. Publieksvragen worden beantwoord.
- ✓ 2 - beperkt. De leerling geeft de bestudeerde thema's aan maar slaagt er niet in deze volledig te beschrijven. Er worden geen conclusies gegeven. Het overbrengen van informatie en de zinsbouw zijn verstaanbaar maar vertonen fouten. Aanduidingen van de voorbereiding en organisatie ontbreken. De visuele hulpmiddelen worden al dan niet vermeld. Publieksvragen worden met beperkte informatie beantwoord.
- ✓ 1 - slecht. De leerling houdt een presentatie zonder aan te geven wat de centrale thema's zijn en welke het belang ervan is. De inhoud is onduidelijk en er worden geen adequate conclusies voorgesteld. De uiteenzetting is moeilijk te volgen. Er is geen indicatie van voorbereiding en organisatie. Publieksvragen worden met zeer beperkte informatie of niet beantwoord.
- ✓ 0 - Er heeft geen presentatie plaatsgevonden.

Het maken van holistische beoordelingen wekt de indruk dat een goede prestatie op de vaardigheid 'oogcontact houden' samengaat met een goede prestatie op vaardigheden 'gegevens verzamelen' en 'selecteren'. Dat klopt uiteraard niet. Hierin schuilt dan ook een belangrijk nadeel van holistische beoordelingen. In de praktijk moet deze kritiek echter genuanceerd worden (zie denkoefening 3.7).

Even stilstaan... 3.7: Holistisch of analytisch?

In vergelijking met de momenteel gangbare beoordelingen zijn de holistische beoordelingen in figuur 3.12 eigenlijk uitermate analytisch.

Kan u akkoord gaan met deze stelling? Waarom wel/niet?

En verder: Welke feedback krijgen leerlingen momenteel na een presentatie? Gaat deze verder dan een allesomvattend cijfer (zoals 6 op 10)? Refereert men hierbij naar specifieke criteria?

Kortom: Zou het gebruik van de bekritiseerbare holistische rubriek al geen grote stap vooruit zijn? Ons inziens wel, vandaar dat wij de stelling onderschrijven.

Om de beschrijving van verschillende types rubrieken af te ronden, stellen we ons even de vraag welk type het meest aangewezen is en de sterkste aanbeveling verdient. Eigenlijk is het onmogelijk om hierover uitspraken te doen; het ene type rubriek is in wezen niet beter of slechter dan het andere. Elk type kan een zinvolle plaats innemen in een assessmentcultuur naargelang het doel van de evaluatie en de daarbij aansluitende voor- en nadelen van het betreffende rubriektype. In figuur 3.13 vatten we deze beide aspecten voor elk rubriektype samen.

Figuur 3.13: Gebruik en voor- en nadelen van de verschillende rubriektypes

WANNEER GEBRUIKEN	RUBRIEKTYPE	VOOR- EN NADELEN
evaluatie van kennis consistentie is zeer belangrijk	taakgebonden	+ zeer betrouwbare evaluatie - tijdsintensief bij het ontwikkelen
evaluatie van vaardigheden bij gebruik van verschillende taken	algemeen	+ inzetbaar bij verschillende taken - weinig specifieke feedback
nood aan gedetailleerde feedback sterkte- zwakteanalyse complexe vaardigheden komen tot zelfevaluatie	analytisch	+ gedetailleerde feedback + consistente scores - tijdsintensief
snel overzicht van de prestaties prestatie vergt slechts één criterium	holistisch	+ snelle beoordeling, overzicht - weinig gedetailleerde informatie - vaak moeilijk in gebruik

3.3.5 Ontwikkel een instrument om leerlingprestaties te beoordelen.

De 5^{de} en laatste stap in het ontwikkelingsproces van criteria en standaarden bestaat uit het ontwikkelen en het gebruiken van instrumenten die het mogelijk maken de criteria effectief te gaan gebruiken. Het ontwikkelen van een betrouwbaar waarderingsproces kan echter ook aangezien worden als een volgende stap in het algemene evaluatieproces dat we hier beschrijven. Herinner u dat we binnen de 3^{de} stap van het evaluatieproces een tussenliggend stappenplan presenteerden (zie fig. 3.6). Het tussenliggend en het algemeen stappenplan kennen vanaf dit punt een zelfde vervolg. In paragraaf 4 van dit deel besteden we verder aandacht aan het vervolg van het evaluatieproces. We beschrijven daar de 4^{de} stap die bestaat uit het ontwikkelen van een betrouwbaar beoordelingsinstrument.

Alles is echter nog niet gezegd over het gebruik van criteria en rubrieken. Om een kwaliteitsvolle evaluatie te garanderen moet deze immers tegemoetkomen aan een aantal kwaliteitseisen. De volgende paragraaf spitst zich daarom verder toe op het beoordelen van criteria en rubrieken.

3.4 Het beoordelen van criteria en rubrieken

Ter afsluiting van de beschrijving van de derde stap uit het evaluatieproces (het ontwikkelen van criteria) rapporteren we een aantal kenmerken van goede criteria of met andere woorden "criteria voor criteria" (Herman et al., 1992; Arter, 1996b). Naast de beschikbaarheid van criteria en rubrieken en het feit dat criteria consistentie moeten kunnen garanderen, kunnen we immers nog andere eisen formuleren ten aanzien van evaluatiecriteria. Indien de criteria niet aan deze eisen voldoen, is het zeer moeilijk om op een consistente, eerlijke en accurate manier de prestaties van leerlingen te beoordelen. Zoals steeds is het belang van deze eisen afhankelijk van de gewichtigheid van de beslissingen die men op basis van de evaluatiegegevens wenst te nemen.

Ongeacht het feit of evaluatiecriteria al dan niet zelf ontwikkeld worden dan wel elders worden ontleend, maakt de volgende set kenmerken het mogelijk de waarde van de criteria te beoordelen. In principe zou elke vraag in de volgende figuur positief beantwoord moeten worden. De kenmerken van criteria die we presenteren, zijn tevens van toepassing op rubrieken, zodat overal waar 'criteria' staat eveneens 'rubrieken' kan gelezen worden. Elk van de aangehaalde kwaliteitskenmerken wordt hieronder verder toegelicht.

Figuur 3.14: Checklist ter beoordeling van criteria

Zijn de evaluatiecriteria ...

- ✓ geschikt om alle belangrijke leeruitkomsten te adresseren?
- ✓ aangepast aan de doelen van de evaluatie?
- ✓ betekenisvol, duidelijk en betrouwbaar?
- ✓ eerlijk en onbevooroordeeld?
- ✓ uitvoerbaar en haalbaar?
- ✓ veralgemeenbaar?

1. Zijn de evaluatiecriteria geschikt om alle belangrijke leeruitkomsten te adresseren?

De criteria moeten alle en tevens enkel de belangrijke leerlingprestaties die men wenst te evalueren, aan de orde stellen. Het belangrijkste criterium ten aanzien van criteria is dan ook het tegemoetkomen aan de validiteitseisen. Dat impliceert dat de rubriek moet reflecteren wat je tijdens het leerproces wil bereiken. Dat uit zich bijvoorbeeld in de vraag of de hoogste score inderdaad datgene omvat wat je vooropgesteld had als leerdoel. De rubriek mag evenwel ook niet meer omvatten dan vooropgesteld. Wanneer de rubriek gebruikt wordt om de aanpak van een wetenschappelijk experiment te evalueren, mogen spelling en grammatica niet van tel zijn. Tenzij men dat wil natuurlijk. Een handige test om te controleren of de cri-

teria en de rubrieken valide beoordelingen mogelijk maken, bestaat in het nagaan hoe de volgende vraag beantwoord moet worden: kan u zich voorstellen dat een leerling die al de vooropgestelde criteria beheerst, niet in staat is goed te presteren op de betreffende taak? Zo ja, dan werkt u met de verkeerde evaluatiecriteria (Wiggins, 1996).

Opmerkelijk is dat aspecten waarover juist de meeste duidelijkheid zou moeten bestaan vaak buiten beschouwing worden gelaten, omdat ze moeilijk kunnen worden uitgeklaard. Dat blijkt uit volgend voorbeeld. Stel dat ik vioollessen volg en van mening ben dat mijn persoonlijke stijl de sterkte is van mijn spel. Wat zou er dan gebeuren als de beoordelaars geen criteria weerhouden in verband met de persoonlijke stijl omdat het moeilijk is die betrouwbaar te scoren? Welke boodschap zou dat impliciet inhouden? En verder: geldt niet hetzelfde voor het beoordelen van creativiteit in oplossingsprocessen en in het schrijven van een essay? Het is echter net belangrijk dergelijke criteria op te nemen, toe te lichten, te communiceren en te bediscussiëren. Uiteraard is het onmogelijk om alle relevant geachte criteria op te nemen. Dat impliceert dat men -indien nodig- moet selecteren en minder belangrijke leeruitkomsten moet weglaten of samenvoegen in een meer omvattend criterium.

2. Zijn de evaluatiecriteria aangepast aan de doelen van de evaluatie?

De criteria moeten het mogelijk maken de eigenlijke doelstellingen van de evaluatie te realiseren. De criteria zullen er bijvoorbeeld anders uitzien naargelang men een holistisch dan wel een analytisch beeld wil van de leerlingprestaties en naargelang de manier waarop men informatie wil verzamelen en rapporteren (kwalitatief of kwantitatief). Wanneer men een niet-taakgebonden rubriek wil ontwerpen, moet men er zich uitdrukkelijk van vergewissen dat de rubriek inderdaad voor een verscheidenheid van evaluatietaken kan worden angewend.

3. Zijn de evaluatiecriteria betekenisvol, duidelijk en betrouwbaar?

De criteria moeten uiteraard betekenisvol zijn voor de verschillende betrokkenen (leerlingen, leerkrachten, ouders, beleids mensen etc.). Je moet criteria opnemen die er toe doen. Stel jezelf daarom de vraag of leerlingen vooruit worden geholpen wanneer ze een rubriek ter beschikking krijgen. Duidelijkheid en betrouwbaarheid zijn vereist om relevante en niet-misleidende resultaten te bekomen. Men spreekt van duidelijke criteria wanneer niet alleen wordt aangegeven waarop je moet letten maar ook hoe hetgeen je observeert er tijdens de evaluatie uitziet. Het volstaat dus niet te benadrukken dat je bijvoorbeeld aandacht moet hebben voor het woordgebruik bij het beoordelen van een opstel. Duidelijke rubrieken illustreren ook wat goed en minder goed woordgebruik inhouden. Iedereen moet verstaan wat een leerling moet presteren om een bepaalde score te behalen en welke kenmerken van de leerlingprestatie maken dat die al dan niet tot een bepaald prestatieniveau behoort. Bij het beschrijven van de karakteristieken voor elk prestatieniveau is het daarom van belang zo weinig mogelijk gebruik te maken van zogenaamd comparatief taalgebruik. Hieronder verstaan we bijvoorbeeld de verdeling in prestatieniveaus op basis van criteria zoals zeer duidelijk, minder duidelijk, minst duidelijk, etc. Met het oog op valide, betrouwbare evaluatie en op leerrijke feedback is het aangewezen voor elk prestatieniveau telkens expliciet de unieke

kenmerken aan te geven. Wat de ene prestatie beter maakt dan de andere moet dus zoveel mogelijk in waarneembare, kwalitatieve termen omschreven worden.

Uiteraard zal het feit dat de criteria voor de beoordelaars duidelijk zijn tevens bijdragen tot een meer betrouwbaar evaluatieproces. De betrouwbaarheid staat of valt immers met de consensus die verschillende beoordelaars kunnen bereiken omtrent de beoordeling van een concrete leerlingprestatie.

4. Maken de evaluatiecriteria een eerlijke en onbevooroordeelde beoordeling mogelijk?

Evenals de evaluatietaken moeten ook de criteria eerlijk zijn. Dit vergt een permanente waakzaamheid voor mogelijke vertekeningen die kunnen optreden bij het definiëren van leeruitkomsten en bij het beoordelen van individuele leerlingenprestaties. Belangrijk is ervoor te zorgen dat de toepassing van criteria niet afhankelijk is van variabelen die door de leerlingen oncontroleerbaar zijn. Criteria hebben pas een diagnostische waarde indien ze uitspraken kunnen doen over door de school te beïnvloeden leerlingenkenmerken. Leerlingen moeten allemaal dezelfde kansen krijgen om een goede score te behalen. De beoordeling mag met andere woorden niet afhankelijk zijn van voor de leerlingen oncontroleerbare variabelen (zoals geslacht, socio-economische achtergrond, voorkennis, beschikbaarheid van het internet thuis, etc.).

5. Zijn de evaluatiecriteria uitvoerbaar en haalbaar?

Herman et al. (1992) rapporteren een aantal redenen waarom het aantal en de complexiteit van dimensies en criteria beperkt moet worden. Allereerst verwijzen ze naar de vaak beperkte mogelijkheden tot intensieve tijd-, werk- en geldinvestering. Hun ervaring is bovendien dat beoordelaars liever niet teveel schalen willen gebruiken om leerlingprestaties te beoordelen. Het evalueren wordt anders te zwaar en dreigt bijgevolg aan betrouwbaarheid in te boeten. Ook voor de leerlingen is het vaak moeilijk om teveel aspecten van hun eigen prestaties tegelijk in ogenschouw te nemen. Het gebruik van de set criteria moet daarom dus ook haalbaar en uitvoerbaar zijn.

Om de stap van het ontwikkelen van criteria af te ronden, willen we u nog twee zaken meegeven. Eerst beschrijven we een mogelijke activiteit die beoogt de kwaliteit van evaluatiecriteria te beoordelen (ontleend aan Rolheiser et al., 2000).

Activiteit 3.2: Beoordelen van criteria

Aan de hand van deze activiteit kan je de voor- en nadelen van specifieke criteria op een rij zetten. De activiteit kan zowel individueel als in groep verlopen. De activiteit is in wezen eenvoudig en bestaat erin voor elk criterium zowel te beschrijven waarom het zeker gebruikt zou moeten worden als waarom het toch geen ideaal criterium is.

Eén van de doelstellingen bij de leerlingen derde graad BSO - Verzorging is: de leerlingen kunnen positieve en negatieve eigenschappen van de eigen persoon als verzorgende opnoemen. Stel dat we criteria zoeken om de reflectie van de leerlingen te evalueren, dan zou de activiteit kunnen leiden tot volgende resultaten.

Mogelijke criteria om de reflectie van leerlingen te evalueren

- Kwaliteit van de reflectie

Zeker gebruiken want: spoort de leerlingen aan de reflectie ernstig te nemen; leert leerlingen dat ze ook kunnen leren van een taak die al af is; stimuleert de leerlingen zichzelf te zien als lerenden. Toch niet ideaal want sommige leerlingen kunnen moeilijkheden ondervinden om uit te drukken wat ze van hun eigen werk vinden.

- **Relevantie**

Zeker gebruiken want: spoort de leerlingen aan hun reflectie te concentreren op zichzelf als verzorgende; stimuleert de leerlingen na te denken over wat belangrijk is voor een goede verzorgster. Toch niet ideaal want de leerlingen kunnen bepaalde interessante reflecties weglaten uit vrees dat ze niet relevant zijn.

- **Netheid en organisatie**

Zeker gebruiken want: stimuleert de leerlingen om de taak zorgvuldig aan te pakken; maakt de evaluatie van de taak voor de beoordelaar(s) eenvoudiger. Toch niet ideaal want er zijn misschien leerlingen die zeer goed kunnen reflecteren zonder hun reflectie te kunnen organiseren. Netheid en organisatie kunnen bovendien een foutieve indruk van kwaliteit geven.

In figuur 3.15 zetten we spelenderwijs een volgende stap in het zoekproces naar criteria voor evaluatiecriteria. Wanneer deze criteria vooropgesteld zijn, is het immers mogelijk een rubriek op te stellen aan de hand van deze criteria. Een rubriek voor het maken van rubrieken dus. In de rubriek worden vier criteria uitgewerkt en wordt een onderverdeling gemaakt in drie kwaliteitsniveaus (rudimentair, werkbaar en nastrevenswaardig). Terloops: maak voor uzelf eens even uit of het hier om een (1) taakgerichte of algemene en (2) analytische of holistische rubriek gaat.

Figuur 3.15: Rubriek voor het evalueren van rubrieken

	rudimentair	werkbaar	nastrevenswaardig
Validiteit	De rubriek omvat niet alle belangrijke criteria of is te omvattend.	De criteria stellen de belangrijkste leeruitkomsten aan de orde, maar moeilijk uit te werken criteria ontbreken.	De criteria maken ook de moeilijke en abstracte leeruitkomsten inzichtelijk. Ze focussen zich op de te beoordelen prestatie.
Duidelijkheid	De uitwerking van de criteria draagt onvoldoende bij tot een consistente score. Aandacht wordt niet gericht.	De meeste criteria maken objectieve en consistente beoordelingen mogelijk. De criteria geven aan waarop de beoordelaar moet letten.	De kwantificeerbare en beschrijvende bewoording leidt tot consistente scores. Duidelijk te onderscheiden prestatieniveaus.
Leerbevorderend	Slaagt er niet in voor het leerproces bruikbare informatie te genereren.	Biedt informatie maar de verwachtingen zijn niet altijd even duidelijk geformuleerd.	Voorziet altijd bruikbare informatie. Duidelijk geformuleerde verwachtingen die de leerlingen uitdagen en stimuleren.
Gebruiks-vriendelijkheid	De rubriek vergt een te grote tijdsinvestering. Of: het aantal prestatieniveaus is te uitgebreid.	De tijdsinvestering is te verantwoorden. De schaal is duidelijk en eenvoudig te hanteren.	Het aantal en de complexiteit van de criteria is aan de doelgroep aangepast. De schaal sluit goed aan bij de doelen van de evaluatie.

4 stap 4: Ontwikkelen van een beoordelingsinstrument

Wanneer voor de te evalueren leerdoelen criteria en rubrieken zijn ontwikkeld, komt het erop aan een betrouwbaar beoordelingsproces op poten te zetten. In deze vierde stap beschrijven we een aantal beoordelingsinstrumenten en hoe die worden samengesteld. Hierbij zal opvallen dat er een zekere overlap bestaat met wat we in vorige stap bespraken. Enerzijds omdat ook rubrieken kunnen gebruikt worden als beoordelingsinstrument en anderzijds omdat elementen zoals het kiezen van een geschikte beoordelingsschaal reeds terloops aangehaald werden. In deze vierde stap worden deze elementen echter verder uitgewerkt.

De mogelijkheden om evaluatiegegevens te verzamelen zijn talrijk en uiteenlopend. Dat bleek toen we de verschillende soorten evaluatietaken bespraken. De evaluatiegegevens die aan de hand van de verschillende taken verzameld kunnen worden, zijn dan ook zeer verscheiden, gaande van keuzes bij een meerkeuzevraag en ingevulde woorden tot essays, presentaties en portfolio's. De beoordeling van deze verschillende vormen van evaluatiegegevens vergt aangepaste instrumenten. In paragraaf 4.1 staan we stil bij de manier waarop we de verzamelde evaluatiegegevens kunnen beoordelen en geven we aan hoe hiertoe geschikte instrumenten ontwikkeld kunnen worden. Hieruit zal blijken dat er centrale rol is weggelegd voor beoordelingschalen. Vandaar dat deze uitdrukkelijk onze aandacht verdienen. Paragraaf 4.2 zet verschillende beoordelingsschalen naast mekaar en gaat in op de keuze van een geschikte schaal.

4.1 Het beoordelen van evaluatiegegevens

Waar staan we in het evaluatieproces? We beschikken over taken die het mogelijk maken evaluatiegegevens te verzamelen. Daarnaast hebben we criteria ontwikkeld die duidelijk maken wat we verwachten van een antwoord, product of prestatie van de leerlingen. Om het onderscheid te maken tussen verschillende prestatieniveaus hebben we de criteria verder uitgewerkt tot rubrieken. Dat brengt ons bij de vraag hoe we, gewapend met de criteria en rubrieken, beoordelingen kunnen uitspreken over de evaluatiegegevens die we aan de hand van de evaluatietaken zullen verzamelen. Het antwoord op deze vraag is zowel afhankelijk van de mate waarin de beoordelaar zelf nog keuzes moet maken als van de aard van de evaluatiegegevens. Dat uit zich in het onderscheid tussen scoringsrichtlijnen en beoordelingsinstrumenten.

4.1.1 Scoringsrichtlijnen

We spreken van scoringsrichtlijnen wanneer de beoordelaar bij het waarderen van evaluatiegegevens zelf geen waardering moet uitspreken over de prestatie van de leerlingen. Er zijn immers vormen van evaluatiegegevens die beoordeeld kunnen worden met een minimale inbreng van de beoordelaar. Hierbij denken we bijvoorbeeld aan de antwoorden op meerkeuzevragen, waarvals vragen, sorteervragen en begrenzend-open-vragen. Leerkrachten spreken over 'corrigeren' wanneer ze de juistheid van de antwoorden op toetsvragen nagaan. Dat dit bij de genoemde

evaluatietaken meestal neerkomt op het toepassen van strikte scoringsrichtlijnen blijkt uit de antwoorden op de volgende toetsvragen.

Figuur 3.16: Scoringsrichtlijnen

Toetsvraag 1. Het jaar 2006 is een schrikkeljaar (juist / fout)?

Antwoord van de leerling: fout.

Scoringsrichtlijn: juist= 0 punten, fout= 1 punt.

Score van deze leerling is dus 1 punt.

Toetsvraag 2. Sorteert de chemische elementen Li, Al, Cs, Fe, Ti, P, S & Cl in volgende chemische families: alkalimetalen, andere metalen en niet-metalen.

Antwoord leerling: alkalimetalen (Li, Cs), andere metalen (Fe, Ti, Al, Cl), niet metalen (P, S).

Scoringsrichtlijn: 4 punten indien alle chemische elementen correct gesorteerd zijn. Trek hiervan per foutief gesorteerd antwoord een half punt af.

Score van deze leerling is dus 3,5 punten (Chloor is geen metaal).

Toetsvraag 3. Wat zijn synoniemen?

Antwoord leerling: Synoniemen zijn woorden die een andere vorm hebben (het zijn verschillende woorden) maar toch hetzelfde betekenen.

Scoringsrichtlijn: 2 punten indien de leerling aangeeft dat synoniemen *ongeveer hetzelfde* betekenen terwijl ze een verschillende vorm hebben, en tevens een correct voorbeeld geeft. 1 punt indien de leerling in eigen woorden aangeeft dat synoniemen *hetzelfde* betekenen terwijl ze een verschillende vorm hebben. 0 punt: andere antwoorden.

Score van deze leerling: 1 punt.

Uit deze voorbeelden blijkt dat leerkrachten bij het scoren van de leerlingenantwoorden zelf bijna geen beslissingen meer moeten nemen over welke score een bepaalde prestatie verdient. In de scoringsrichtlijnen worden hierover duidelijke instructies gegeven. Met de richtlijnen in de hand zal een leek tot dezelfde score komen als een ervaren leerkracht. Het expliciet uitwerken van scoringsrichtlijnen is echter geen gangbare praktijk in de huidige evaluatiecultuur. Leerkrachten hanteren wel richtlijnen maar deze blijven veelal onuitgesproken. Dit zou er toe kunnen leiden dat niet altijd dezelfde richtlijnen gebruikt worden. Op verschillende manieren dreigt de betrouwbaarheid van de evaluatie daardoor in het gedrang te komen. De meest aangehaalde meetfouten zijn: naarmate de correctieperiode vordert, wordt steeds minder nauwkeurig te werk gegaan; de score is afhankelijk van de kwaliteit van het voorgaande antwoord dat verbeterd werd; niet relevante kenmerken van het antwoord worden in rekening gebracht (schrijfstijl, lengte van het antwoord); vooroordelen en de stemming van de beoordelaar beïnvloeden de score (Standaert & Troch, 1998). Zonder expliciete scoringsrichtlijnen kan men dus geen kwaliteitsvolle evaluatie garanderen.

Scoringsrichtlijnen zullen er naargelang de gebruikte evaluatietaak telkens anders uitzien. In wezen komt het er op neer dat scoringsrichtlijnen steeds uitgebreider en complexer worden naarmate de subjectieve inbreng van de beoordelaar groter wordt. Meerkeuzevragen en waarvals vragen behoeven amper richtlijnen, het antwoord is of juist of fout. Begrenzende-openvragen vergen wel richtlijnen. Deze bestaan meestal uit een concreet antwoordmodel waarop

volgens een afgesproken schaal de score gebaseerd wordt. De evaluatie van een essay vereist dan weer een hele rubriek. Dit brengt het onderscheid tussen scoringsrichtlijnen en beoordelingsinstrumenten aan de orde. Scoringsrichtlijnen kunnen de objectiviteit waarborgen omdat beoordelaars tijdens de correctie zelf geen beoordeling moeten uitspreken. Dit kan echter maar tot op een zekere hoogte. Denken we bijvoorbeeld aan het criterium 'goed oogcontact onderhouden' bij het evalueren van een mondelinge presentatie. Een mogelijke scoringsrichtlijn zou dan kunnen zijn: 2 punten als de leerling meer dan viermaal per minuut oogcontact heeft met het publiek, etc. Dergelijke scoringsrichtlijnen hebben echter geen zin en zouden in het geheel onmogelijk te gebruiken zijn (zoek maar eens een scoringsrichtlijn voor lichaamshouding). Om dergelijke elementen te evalueren kan men niet voorbij aan de inbreng van de individuele beoordelaar. Wanneer beoordelaars wel een beoordeling moeten uitspreken, (en dus niet alleen vooropgestelde richtlijnen moeten toepassen) kan er gebruik gemaakt worden van zogenaamde beoordelingsinstrumenten (zie paragraaf 4.1.2). Dat is bijvoorbeeld het geval bij het beoordelen van producten en bij het observeren van authentieke leerlingenprestaties. Kortom, bij het evalueren van (authentieke) vaardigheidstaken die wij doorheen dit handboek bepleiten.

Bij het scoren of corrigeren van evaluatiegegevens moeten we het volgende duidelijk benadrukken. Een evaluatiescore is nog geen beoordeling. Na het corrigeren weten we bijvoorbeeld dat de leerling in kwestie 26 van de 40 goed/foutvragen correct beantwoord heeft. Al te vaak aanzien leerkrachten deze score tevens als beoordeling. Zo wordt er bijvoorbeeld van uitgegaan dat voldoende betekent dat men 50% of 60% van de punten behaalt. De helft halen staat dus gelijk met een voldoende. Wie de helft van de opgaven juist heeft krijgt dan een 5 op 10; wie 26 van de 40 vragen juist heeft, krijgt een 6,5. Hierbij wordt het fundamentele onderscheid tussen scores en beoordelingen over het hoofd gezien. In sommige gevallen kan uit het goed oplossen van de helft van de opgaven blijken dat de leerling een slechte prestatie heeft neergezet, in een andere situatie geeft dezelfde score blijk van een puike prestatie. Om van een evaluatiescore tot een beoordeling te komen, is er een referentiekader nodig. Dat kan bijvoorbeeld bestaan uit een aantal richtlijnen die aangeven wat je van de leerlingen verwacht. Zo kan het correct oplossen van minstens 4 op 5 vragen bijvoorbeeld de beoordeling 'goed' verdienen terwijl het correct oplossen van de helft tot 3 op 5 vragen de beoordeling 'voldoende met leemtes' krijgt. Op basis van de leerdoelen die vooropgesteld werden, volstaat het correct oplossen van de helft van de opdrachten in dat geval dus niet om van een 'voldoende' te kunnen spreken.

4.1.2 Beoordelingsinstrumenten

Terwijl scoringsrichtlijnen de betrouwbaarheid van een beoordeling trachten te garanderen, kunnen beoordelingsinstrumenten niet meer doen dan de betrouwbaarheid optimaliseren. Een perfecte betrouwbaarheid mag dan wel onrealistisch zijn, toch streven we ze na. Vandaar de nood aan evaluatiecriteria, vandaar de nood aan instrumenten om verzamelde of geobserveerde evaluatiegegevens te beoordelen en vandaar de nood om het beoordelingsproces ernstig te nemen (zie stap 5). In de praktijk van alternatieve evaluatievormen zijn twee beoordelingsinstru-

menten gebruikelijk: evaluatielijsten en rubrieken. Beide types instrumenten worden in de volgende paragrafen toegelicht.

1. Evaluatielijsten als beoordelingsinstrumenten

Een evaluatielijst heeft als doel de beoordelaars van een leerlingprestatie bij te staan in het beantwoorden van de vraag wat de kwaliteitsaspecten in een bepaald product of in een bepaalde leerlingprestatie zijn (ASCD, 1996). Evaluatielijsten zijn instrumenten om leerlingprestaties te beoordelen en sluiten nauw aan bij de rubrieken en de criteria die reeds werden ontwikkeld voor de vooropgestelde evaluatietaak. Hoewel ze op mekaar kunnen lijken, zijn rubrieken dus niet helemaal hetzelfde als evaluatielijsten. Evaluatielijsten vloeien als het ware voort uit rubrieken. In tegenstelling tot rubrieken zijn evaluatielijsten steeds analytisch van aard. Ze worden ontworpen om de prestaties van leerlingen bij concrete taken te beoordelen. Bijvoorbeeld: is één van de criteria in de rubriek 'een goed gebruik van verschillende media', dan wordt dat in een evaluatielijst vertaald in 'de leerling maakt gepast gebruik van transparanten en videobeelden'. Rubrieken zijn meestal bedoeld om algemene – niet taakgebonden – uitspraken te doen (in het geval van algemene, holistische rubrieken). Om een algemene score van een leerling op een rubriek te verantwoorden, kan men bijvoorbeeld verwijzen naar de scores op een evaluatielijst (ASCD, 1996).

Figuur 3.17 geeft een illustratie van een mogelijke evaluatielijst. De evaluatielijst maakt een onderscheid tussen vier prestatieniveaus: (1) werk aan de winkel, (2) aanvaardbaar, (3) goed, (4) een voorbeeldprestatie.

Figuur 3.17: Een evaluatielijst voor het beoordelen van groepswork

	1	2	3	4
De leerling komt goed voorbereid naar het groepswork	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling brengt individuele opdrachten tijdig tot een goed einde	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling participeert op een constructieve wijze aan het groepswork	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling stimuleert anderen om constructief mee te werken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling luistert goed naar de inbreng van anderen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling beargumenteert het eigen standpunt op duidelijke en doordachte manier	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling laat op een aangename manier merken dat hij/zij het met iets oneens is	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling kan compromissen bereiken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling helpt de groep om richtlijnen te volgen en het tijdschema te respecteren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling bewaakt de positieve relaties binnen de groep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerlingen herkennen problemen en kunnen er gepast op reageren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling neemt bij beslissingen alle relevante informatie in acht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling wisselt duidelijk en accuraat informatie uit met de andere groepsleden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

De samenstelling van evaluatielijsten is in wezen eenvoudig. De meeste evaluatielijsten bevatten een aantal uitspraken over kwaliteitsaspecten van een leerlingprestatie waarover de beoor-

delaar met behulp van een schaal een beoordeling uitspreekt. Een uitspraak zouden we kunnen omschrijven als een statement over een specifiek kwaliteitsaspect (een criterium) zoals (1) de presentatie is duidelijk gestructureerd of (2) de lijnen van de grafiek maken goed gebruik van de beschikbare ruimte. Evaluatielijsten zijn dus een opsomming van relevante criteria waarover aan de hand van een gepaste evaluatieschaal uitspraken worden gedaan. Het ontwikkelen van criteria werd hoger in paragraaf 3.3 reeds besproken, terwijl schalen in paragraaf 4.2 aan bod komen. Na het beschrijven van de verschillende soorten schalen die gebruikt kunnen worden in beoordelingsinstrumenten, gaan we aldaar eveneens in op de keuze van een geschikte schaal.

Evaluatielijsten kunnen zo ontworpen worden dat ze de mogelijkheid bieden differentiërend tewerk te gaan. Men kan de leerkracht in een eerste kolom van de evaluatielijst de kans geven het relatieve belang van het betreffende criterium aan te geven (eventueel in samenspraak met de leerlingen). Bij belangrijk geachte uitspraken (criteria) kunnen dan zo nodig meer punten verdeeld worden dan bij minder belangrijk geachte uitspraken. Bij de ene leerling kan aan een bepaald criterium meer belang worden toegekend dan bij een andere. Het belang van de verschillende uitspraken kan dus variëren. Op die manier is het mogelijk om een duidelijke nadruk te leggen op bepaalde dimensies van de taak zodat het mogelijk wordt om de evaluatielijsten te individualiseren, om te differentiëren voor leerlingen met specifieke behoeftes en om het ontwikkelingsproces van de leerlingen gericht te volgen en te sturen.

Figuur 3.18: Differentieerbare evaluatielijst

	Mogelijke evaluatiepunten	Zelfevaluatie
De titel van de grafiek geeft duidelijk het verband aan tussen de assen.	10	_____
Er werd een geschikt grafiektype gebruikt.	20	_____
De assen dragen een duidelijk label.	10	_____
De lijnen of balken gebruiken de beschikbare ruimte goed.	30	_____
....	...	_____
De grafiek is makkelijk te interpreteren.	20	_____
Totaal	100	_____

Evaluatielijsten hoeven niet uitsluitend door leerkrachten gebruikt te worden, ook leerlingen kunnen ermee aan de slag. In dat geval wordt er in de lijst een ruimte voorzien voor leerlingen om zichzelf te beoordelen en eventueel ook voor leerlingen om elkaar te evalueren. Afhankelijk van het ontwikkelingsniveau zullen de evaluatielijsten en de gebruikte schalen er anders uitzien. In het lager onderwijs hebben evaluatielijsten bijvoorbeeld vaak de vorm van een aantal vragen die aan leerkrachten en/of leerlingen gesteld worden en die uitnodigen om evaluatieve uitspraken te doen. Meestal worden kwalitatieve schalen gebruikt (zoals *buitengewoon, ok, nog aan werken*). Hoewel de meeste evaluatielijsten worden ontwikkeld door leerkrachten of groepen leerkrachten is het bovendien goed mogelijk dat ook leerlingen aan het ontwikkelingsproces

participeren. Het maken van evaluatielijsten kan voor hen immers een waardevolle leerervaring zijn.

2. Rubrieken als beoordelingsinstrumenten

We stelden hoger dat evaluatielijsten een uitwerking zijn van rubrieken. Rubrieken hoeven evenwel niet altijd in evaluatielijsten te worden omgezet om ze te gebruiken als beoordelingsinstrument. Sommige rubriektypes kunnen immers gebruikt worden om analytische uitspraken te doen, andere weer niet. Dat laatste geldt voor de zogenaamde holistische rubrieken. Als je een leerlingprestatie wil beoordelen aan de hand van een holistische rubriek moet je eerst meer gedetailleerde informatie verzamelen omtrent het al dan niet tegemoetkomen van de prestatie(s) aan de vooropgestelde criteria. Dat kan aan de hand van evaluatielijsten of aan de hand van analytische rubrieken. Het gebruik van een holistische rubriek impliceert dat beoordelingen noodgedwongen algemeen blijven (bekijk voor alle duidelijkheid eerst het voorbeeld van een holistische rubriek nogmaals (paragraaf 3.3.4 punt 2). Holistische rubrieken maken dus niet in dezelfde mate duidelijk op welke criteria (uitspraken) de beoordeelde nu al dan niet goed scoorde. Het komt er op aan een zo getrouw mogelijk prestatieniveau te kiezen. Een dergelijke beoordeling mag dan wel minder informatief en minder gedetailleerd zijn, dat is geen probleem wanneer men algemene beoordelingen wil uitspreken.

Hieruit onthouden we dat holistische rubrieken pas gebruikt kunnen worden om beoordelingen uit te spreken wanneer de beoordelaar reeds over gegevens beschikt afkomstig van evaluatielijsten of analytische rubrieken. Bij analytische rubrieken liggen de zaken dan ook anders. Deze rubrieken kunnen wel rechtstreeks als beoordelingsinstrumenten gebruikt worden. Ze tonen immers duidelijk aan hoe een leerlingprestatie die tot een bepaald kwaliteitsniveau behoort, er dient uit zien. Die kwaliteitsindicaties ondersteunen de beoordelaar door voor elk criterium aan te geven wat de verschillende karakteristieken zijn van elk prestatieniveau. Aan de hand van de beschrijvingen in de rubriek kan de beoordelaar uitmaken tot welk niveau de leerlingprestatie behoort.

Bij het bespreken van een aantal alternatieve evaluatievormen in deel 4 zal uit de illustraties duidelijk worden op welke manier evaluatielijsten en rubrieken gebruikt kunnen worden als beoordelingsinstrumenten.

4.2 Over beoordelingsschalen

Aan de hand van schalen doen beoordelaars uitspraken over elk van de uitspraken in de evaluatielijst of de analytische rubriek. Beoordelingsschalen zijn dan ook een essentieel onderdeel van evaluatielijsten en rubrieken. In de volgende paragrafen besteden we daarom aandacht aan de verschillende soorten schalen en aan de keuze van een geschikte schaal.

4.2.1 De soorten schalen

Een schaal is een reeks vaststellingen of beoordelingen (getallen of beschrijvingen) die volgens een bepaald voorschrift gekoppeld worden aan de prestaties van leerlingen die men tijdens het evolutieproces heeft waargenomen (Thio, 1988). Er zijn verschillende types schalen die men kan hanteren bij het evalueren van leerlingprestaties. Het voornaamste onderscheid tussen deze verschillende schalen gaat terug op de vraag of de schaal zich beperkt tot het weergeven van observaties dan wel of de schaal beoordelingen uitspreekt. Vandaar de onderverdeling in observatie- en beoordelingsschalen.

1. Observatieschalen: checklists

Een specifieke soort schaal zijn de observatieschalen of checklists. Bij het gebruik van observatieschalen beantwoordt men enkel de vraag of een gedraging zich al dan niet heeft voorgedaan, zonder te beoordelen hoe goed elk van de gedragingen uitgevoerd werd. De schaal beperkt zich tot de beoordelingen ja - nee, aanwezig - afwezig, volbracht - niet volbracht en dergelijke. In deze gevallen is de evaluatielijst een voorbeeld van een checklist. Een checklist is dus een lijst van dimensies, karakteristieken of gedragingen (de zogenaamde uitspraken) die in essentie worden gescoord met ja of nee. Het is een observatie-instrument waarmee wordt aangegeven of een leerling bepaalde kennis, vaardigheden of houdingen al dan niet bezit (Herman et al, 1992; Thio, 1988). Een checklist is dus in principe geen beoordelingsschaal omdat er aan de geobserveerde gedragingen geen waardeoordeel wordt gekoppeld.

Een checklist over het respecteren van veiligheidsvoorschriften bij het zagen van houten planken ziet er bijvoorbeeld als volgt uit. De kolom NVT (niet van toepassing) werd toegevoegd om aan te geven dat de leerling tijdens de observatieperiode niet de kans heeft gehad om een bepaald criterium te demonstreren.

	JA	NEE	NVT
De leerling concentreert zich wanneer hij de zaagmachine hanteert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling zet de machine af tussen twee werkmomenten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling draagt aangepast schoeisel en handschoenen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling vraagt hulp wanneer te grote stukken moeten bewerkt worden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De leerling controleert regelmatig de staat van de zaagmachine.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Totaal			

Een verder uitgewerkte vorm van observatieschalen zijn de zogenaamde descriptieve schalen. Deze schalen gaan verder dan het scoren in de vorm van een 'ja-dan-nee'-schaal. Descriptieve schalen bestempelen de leerlingprestaties op basis van neutrale termen zonder te verwijzen naar

een onderliggende standaard. Voorbeelden van een descriptieve schaal zijn: (1) geen poging gedaan de taak te volbrengen...taak gedeeltelijk volbracht...taak volbracht...taak meer dan volbracht en (2) niet met de taak bezig...poging tot aanpakken taak...minimale aandacht voor de taak...taak aangepakt maar niet uitgewerkt...taak uitgewerkt.

2. Beoordelingsschalen

Schalen kunnen echter ook onmiddellijk een beoordeling koppelen aan wat werd geobserveerd. In plaats van enkel iets vast te stellen, spreekt men eveneens een waardering uit over wat vastgesteld werd. Dat kan aan de hand van grafische, numerieke en kwalitatieve beoordelingsschalen of een combinatie daarvan.

a. Grafische schalen

Een grafische schaal bestaat meestal uit een continuüm met aan de uiteinden de meest negatieve en de meest positieve graad van een bepaalde gedraging of prestatie (zoals routinematig vs. origineel; slordig en onverzorgd vs. zeer zorgzaam en ordelijk; goed gestructureerd vs. oeverloos). Het is aan de beoordelaar om op de gepaste plaats op het continuüm de gewenste beoordeling aan te duiden. Bij het criterium nauwkeurigheid zou de grafische schaal er bijvoorbeeld als volgt kunnen uitzien:

□	
Veel fouten, niet bekommerd om accuratesse	Zelden of geen fouten, steeds bekommerd om accuraat werk

Grafische schalen dragen zelden bij tot de objectiviteit van een beoordeling. De verschillende posities op de schaal zijn immers niet omschreven zodat de beoordeling wordt overgelaten aan de interpretatie van diegene die de schaal invult en diegene die de schaal leest.

b. Numerieke schalen

Een numerieke schaal gebruikt getallen of bepaalde codes (lettercodes) als een continuüm van prestatieniveaus. De evaluator geeft dus een waardecijfer of -letter aan het geobserveerde gedrag. Te denken valt bijvoorbeeld aan de alom bekende 10-puntenschaal of aan het gradensysteem zoals we dat uit de Verenigde Staten kennen (A+, A, B, C, D, ...). Het aantal onderdelen op de schaal hangt veelal samen met het aantal prestatieniveaus dat in de rubriek werd onderscheiden. Numerieke schalen bestaan dan ook vaak uit 4, 5 of 6 onderverdelingen. Ook bij deze schalen geldt dat niet noodzakelijk iedereen hetzelfde verstaat onder een bepaalde score.

Totale betrouwbaarheid mag dan wel uitgesloten zijn, de vorige twee schaaltypes zijn in mindere mate geschikt om de betrouwbaarheid zo groot mogelijk te maken.

c. Kwalitatieve schalen

Een kwalitatieve schaal gebruikt adjectieven in plaats van getallen of letters om de prestaties van leerlingen te karakteriseren. Meestal bestaan de verschillende schaalni-

veaus uit verbale omschrijvingen van de verschillende kwaliteitsniveaus van de te evalueren prestaties. Voorbeelden zijn:

- ✓ uitstekend...goed...behoeft verbetering
- ✓ voorbeeldig antwoord... competent antwoord... kleine tekorten maar voldoende... ernstige tekorten maar bijna voldoende... probleem aangevat maar inadequaat antwoord... niet in staat probleem aan te pakken... geen poging
- ✓ vakmanschap...leerjongen...beginneling

Ook het gebruik van symbolen in de vorm van gezichten met een grote, kleine of neutrale glimlach zou als een vorm van een kwalitatieve schaal aangezien kunnen worden.



d. Gecombineerde schalen

Uit de vorige schaaltypes blijkt dat schalen kunnen worden weergegeven in de vorm van cijfers, letters of woorden. Ook combinaties tussen letters of cijfers en kwalitatieve beschrijvingen zijn echter mogelijk. Numerieke schalen mogen dan wel de mogelijkheid bieden om op te tellen en te middelen, ze zijn eerder moeilijk te interpreteren als men geen zicht heeft op de achterliggende criteria. Verschillende mensen zullen bijvoorbeeld een score van 6 op 10 verschillend interpreteren. Om dergelijke problemen te vermijden, kan men een combinatie van numerieke en kwalitatieve schalen hanteren. Zo kan men een numerieke score laten vergezellen van een evaluatie gaande van inadequaat tot competent. Ook aan cijfers kan dus een beoordelende interpretatie worden gekoppeld. In het geval van de traditionele 10-puntenschaal kan dit er bijvoorbeeld als volgt uitzien (De Groot et al., 1983).

Figuur 3.19: Een gecombineerde 10-puntenschaal

10 – uitstekend	5 – amper voldoende
9 – zeer goed	4 – onvoldoende
8 – goed	3 – zeer onvoldoende
7 – ruim voldoende	2 – slecht
6 – gewoon voldoende	1 – zeer slecht

Sluit deze verdeling aan bij uw eigen intuïtieve invulling van een 10-puntenschaal? Wellicht komen deze beoordelingen grotendeels overeen met de betekenis die leerkrachten, leerlingen en ouders aan de verschillende punten geven. Grotendeels, want het combineren van de gangbare 10-puntenschaal met een kwalitatieve schaal is niet zonder problemen. De koppeling tussen beide schalen blijft meestal impliciet en de kans is

daarom groot dat er te weinig overeenstemming bestaat omtrent welke kwalitatieve beschrijving verschillende mensen koppelen aan een bepaalde score. Bovendien kan men zich de vraag stellen of het wel zinvol is een onderscheid te maken tussen 10 prestatieniveaus. Hoe informatief is het onderscheid tussen een 7 en een 8, of tussen een 4 en een 3? Terloops willen we opmerken dat naarmate men meer categorieën wenst te gebruiken de kans om niet-consistente beoordelingen te maken groter wordt (en dus de betrouwbaarheid lager).

De keuze van het type schaal en het aantal prestatieniveaus is des te belangrijker omdat ze bepalend is voor de wijze waarop en de hoedanigheid waarin evaluatiegegevens later kunnen worden teruggekoppeld en gerapporteerd. Het selecteren van een geschikte schaal moet dus goed doordacht gebeuren. In de volgende paragraaf zetten we daarom de voor- en nadelen van een aantal schaaltypen naast mekaar.

4.2.2 De keuze van een geschikte schaal

Een aantal problemen met numerieke scores maken het aangewezen bij het beoordelen van leerlingenprestaties kwalitatieve of gecombineerde beoordelingsschalen te gebruiken. De volgende paragrafen beargumenteren kort onze terughoudendheid ten aanzien van numerieke scores en het pleidooi voor alternatieven.

Numerieke beoordelingsschalen zijn zeer sterk ingeburgerd in het Vlaamse onderwijs en dit meestal in de vorm van een 10-puntenschaal. Leerlingen krijgen op verschillende toetsen punten die dan worden verwerkt tot een eindscore per vak, meestal door te middelen. Dat gemiddelde wordt doorgaans uitgedrukt op noemer 10 of noemer 20 (resp. 7/10 of 14/20). Vervolgens wordt vaak in een percentage een vakoverschrijdend eindoordeel gerapporteerd. Dit mondt uit in globale eindbeoordelingen aan het einde van het schooljaar in de trend van: 'Peter heeft slechts 52% terwijl Ellen 87% behaalde'. Dergelijke praktijken om de evaluatiegegevens te communiceren zijn alom bekend. Toch is het gebruik van deze numerieke scores niet vrij van kritiek (Standaert & Troch, 1999; Janssens et al, 2000; Linn, 1996). Volgende opsomming zet de belangrijkste kritieken op een rij.

- ✓ Wat betekent het gegeven dat Sven een 6 heeft gekregen voor de afgelopen stageperiode. En wat met Ann die een uitslag van 20 op 30 behaalt op haar leestoets? Mag zij met dat resultaat tevreden zijn? Zonder een goede interpretatie is de betekenis van deze punten allermindst duidelijk en informatief. Numerieke scores hebben daarom duidelijke beperkingen bij het begeleiden van leerlingen en bij het verzorgen van feedback. Hetzelfde geldt overigens voor het gebruik van letters zonder meer.
- ✓ De cijfers hebben in de praktijk telkens betrekking op het resultaat van het leerproces en niet op de manier waarop het resultaat tot stand is gekomen. Ook dit maakt de informatieve waarde van cijfers en letters beperkt.

- ✓ Het optellen en middelen van numerieke scores maakt dat belangrijke evoluties in het leerproces uit het oog worden verloren. Een leerling met respectievelijk een 2, 3, 7 en 8 zal gemiddeld slechts een 5 behalen hoewel zijn laatste prestaties aantonen dat hij de verwachte kennis en vaardigheden onder de knie heeft. De beoordeling is overigens dezelfde voor de leerling die achtereenvolgens een 8, 7, 3 en 2 behaalde. De feedback die beide leerlingen krijgen, zou echter fundamenteel moeten verschillen.
- ✓ Hoewel punten volgens velen (onderwijsgevers en ouders) de leerlingen kunnen motiveren, hebben ze het grote nadeel dat ze onderlinge vergelijking en competitie uitlokken. Ook al was competitie inherent aan de traditionele onderwijs- en evaluatiebenadering, vandaag wordt aangenomen dat men het competitieve aspect bij elke vorm van evaluatie moet vermijden. En daarom: is het zinvol steeds een onderscheid te maken tussen een 9 en een 10 terwijl men beide leerlingen zou kunnen beoordelen met de uitspraak 'een excellente prestatie'?
- ✓ Numerieke scores zetten leerlingen aan te studeren om (goede) punten te behalen en leiden zo de aandacht af van de onderwijsdoelen die echt belangrijk zijn. Ten gevolge van het gangbare puntensysteem dreigen leerlingen zich enkel te zullen interesseren voor het behalen van hoge cijfers en niet zozeer voor het leren op zich. Er doet zich een verschuiving voor van een taakgerichtheid naar een prestatiegerichtheid.
- ✓ De beoordelingsstrategie die gebaseerd is op het toekennen van cijferscores is niet geschikt voor de functies en doelen van evaluatie in een assessmentcultuur. Cijferscores passen dan ook niet bij het gebruik van alternatieve evaluatievormen. Wat lijkt u het meest waardevol: de beoordeling in de vorm van een 6 op 10 of de beoordeling voldoende in een analytische rubriek waarbij voor elk criterium wordt aangegeven waaraan de leerling nog moet werken om een betere beoordeling te verdienen?

Cijfers oogsten dus verschillende kritieken. De belangrijkste kritiek op het gebruik van cijferscores is het feit dat numerieke schalen in principe voor niemand duidelijk zijn zolang men er geen duidelijke interpretatie aan toekent (Standaert & Troch, 1998). Dat betekent dat het zinvol is om aan cijfers een woordelijke beoordeling te verbinden, om alzo te komen tot een zogenaamd beoordelingscijfer. Een beoordelingscijfer (een kwalitatief cijfer) drukt symbolisch uit of de score al dan niet overeenstemt met een bevredigende prestatie. Achter het beoordelingscijfer gaat een expliciete woordelijke beoordeling schuil (vooreen voorbeeld, zie figuur 3.20). Wanneer leerlingen en leerkrachten voldoende vertrouwd zijn met de gebruikte schaal hoeft de kwalitatieve beoordeling niet voortdurend meer te worden vermeld. Zolang de cijfers maar geen eigen leven gaan leiden en de kwalitatieve beoordeling, indien nodig, steeds opnieuw ter hand kan worden genomen.

Uit de opmerkingen bij de traditionele 10-puntenschaal leren we twee belangrijke zaken: (1) beperk de beoordeling niet tot puntscores en (2) beperkt het aantal categorieën tot een werkbaar aantal. Dat maakt dat goede kwalitatieve beoordelingsschalen er bijvoorbeeld als volgt kunnen uitzien.

- ✓ In de meeste evaluaties volstaat het een onderscheid te maken tussen vier categorieën. Of men deze weergeeft met 1, 2, 3, 4 of A, B, C, D maakt in principe niet veel uit (op het gegeven na dat letters de garantie bieden dat ze niet kunnen worden opgeteld en gemiddeld). Belangrijk is wel dat deze letters of cijfers steeds vergezeld worden van een kwalitatieve beoordeling (een interpretatie). Het zijn dus beoordelingsletters of beoordelingscijfers. Viergradenschalen zullen er aldus als volgt uitzien.

Figuur 3.20: Voorbeelden van viergradenschalen

- ✓ A (of 1) – zeer goed
- ✓ B (of 2) – goed
- ✓ C (of 3) – voldoende
- ✓ D (of 4) – onvoldoende
- ✓ 4 – overtreft de verwachtingen
- ✓ 3 – komt tegemoet aan de verwachtingen
- ✓ 2 – benadert de verwachtingen
- ✓ 1 – er wordt meer verwacht
- ✓ Goed: indien de vaardigheid aanwezig is en optimaal wordt gedemonstreerd (G of 1)
- ✓ Voldoende: indien de vaardigheid aanwezig is en correct wordt gedemonstreerd (V of 2)
- ✓ Voldoende met leemten: indien de vaardigheid aanwezig is maar gedemonstreerd wordt met tekorten (VL of 3)
- ✓ Onvoldoende: indien de vaardigheid in onvoldoende mate aanwezig is (O of 4)

Elk van deze beoordelingsschalen benadrukt een aantal ontwikkelingsstadia. De gekozen categorieën resulteren echter telkens in een andere verdeling. Hierbij wordt doelbewust vermeden dat er teveel negatieve categorieën zouden zijn. Bij de kwalitatieve omschrijving van de categorieën tracht men immers pejoratief taalgebruik te vermijden (Chapman, 1990). Er wordt eerder gesproken over 'onvoldoende', 'pover', 'behoeft nog werk' en 'nog niet volledig ontwikkeld' dan kortweg over 'zwak' en 'slecht'.

- ✓ Een interessante schaal wordt gebruikt in de zogenaamde '6-punten, 2-beslissingen'-rubriek (zie figuren 10 en 21). Deze structuur van rubrieken is interessant omdat men als beoordelaar slechts twee beslissingen moet nemen om tot een eindoordeel te komen. Met twee beslissingen wordt de prestatie dus op een kwaliteitsschaal met 6 onderverdelingen geplaatst. In de eerste beslissing moet worden uitgemaakt of de prestatie voortreffelijk dan wel pover is. Vervolgens moet nog worden uitgemaakt of de prestatie beter, overeenstemmend dan wel slechter was dan het oordeel uit de eerste beslissing. De kwalitatieve beoordelingsschaal bij deze rubrieken ziet er bijvoorbeeld als volgt uit.

Figuur 3.21: Voorbeeld van de '6-punten, twee-beslissingen'-schaal

- A – Buitengewoon
- B – Voortreffelijk
- C – Voortreffelijk, ware het niet dat één of twee belangrijke elementen niet voortreffelijk zijn

- D – Pover, ware het niet dat één of twee belangrijke elementen goed zijn
- E – Pover
- F – Zeer slecht

- ✓ In het geval van kleine procesevaluaties kan het aantal categorieën verder beperkt worden tot twee. Een schaal met de onderverdeling V = voldoende en O = onvoldoende kan in dat geval volstaan.

Alvorens dit deel over de keuze van een geschikte schaal af te ronden, willen we benadrukken dat -omwille van de beperkte waarde van een cijfer en een letter- een kwalitatieve beoordeling meer aangewezen is. De gedetailleerde en bruikbare informatie die men leerlingen kan bezorgen aan de hand van een woordelijke uitleg maakt dat leerlingen kunnen leren uit hun fouten en eventuele hiaten en tekorten kunnen wegwerken. Bij mindere prestaties kan bovendien toch gezocht worden naar het positieve en stimulerende. Bij de 6^{de} stap in het evaluatieproces (aangaande de communicatie en het gebruik van evaluatiegegevens) komt de rapportage van evaluatiegegevens verder aan bod. Om het voorgaande af te sluiten, willen we echter eerst nog even stilstaan bij het hardnekkig voortbestaan van puntensystemen.

Even stilstaan... 3.8: Het hardnekkig voortbestaan van puntensystemen

We hebben reeds opgemerkt dat een cruciale factor in de kanteling van een toets- naar een assessmentcultuur is weggelegd voor de gangbare onderwijs- en evaluatievisies van leerkrachten. Breng even een vertrouwde onderwijscontext voor de geest en ga eens na of de leerkrachten in die context snel geneigd zullen zijn om bovenstaand pleidooi voor het afwijken van puntensystemen te ondersteunen. De argumenten van voorstanders zullen wellicht overeenstemmen met het bovenstaande. Ga daarom eens na welke argumenten tegenstanders zouden aanhalen om zich halsstarrig te blijven vastklampen aan het bestaande puntensysteem. Ondersteunt u deze argumenten?

Wanneer je met zekerheid beschikt over een voldoende betrouwbaar beoordelingsinstrument kan je overgaan tot de eigenlijke verzameling en beoordeling van evaluatiegegevens van de leerlingen. Hoe goed de criteria en rubrieken ook ontwikkeld zijn, de garantie voor een kwaliteitsvolle evaluatie is er pas wanneer deze instrumenten ook op een gepaste manier gebruikt worden.

5 Stap 5: Verzamelen en beoordelen van evaluatiegegevens

Het feit dat alternatieve evaluatievormen in belangrijke mate een beroep doen op de menselijke perceptie en interpretatie maakt dat aanzienlijke aandacht moet worden besteed aan de kwaliteit van zowel het verzamelen als het beoordelen van evaluatiegegevens. Deze bezorgdheid staat centraal in de vijfde stap van het evaluatieproces. Terwijl er in de vorige stappen sprake was van een zorgvuldige voorbereiding van verschillende evaluatieaspecten, gaan we hier effectief aan

de slag met de ontwikkelde taken en beoordelingsinstrumenten. Hoe zorgvuldig de voorbije stappen ook doorlopen werden en hoe kwaliteitsvol de ontwikkelde taken en instrumenten ook mogen zijn, indien deze op een ondoordachte manier gebruikt worden, kan er geen sprake zijn van een kwaliteitsvolle evaluatie.

In de volgende paragraaf staan we eerst stil bij een aantal relevante aspecten van een kwaliteitsvolle gegevensverzameling en -beoordeling. Dat zal neerkomen op een korte opfrissing van wat we in deel 2 bespraken toen we op zoek gingen naar de kenmerken van een goede evaluatie. Daarna zetten we een aantal elementen op een rij die het verschil kunnen maken tussen een kwaliteitsvolle en een bekritiseerbare evaluatie. Ook de vraag door wie de gegevens verzameld worden en door wie de ontwikkelde instrumenten gebruikt worden, zou binnen deze stap moeten ingevuld worden. Om overlap te vermijden, worden deze aspecten echter beschreven in deel 4 wanneer we onze aandacht richten op de betrokkenheid van leerlingen als een kenmerk van alternatieve evaluatievormen.

5.1 Wanneer spreekt men van een kwaliteitsvol beoordelingsproces?

In deel 2 presenteerden we een aantal kwaliteitseisen die kunnen verzekeren dat men correcte uitspraken doet op basis van de evaluatiegegevens. Omdat op basis van de resultaten van een evaluatie vaak belangrijke beslissingen worden genomen, is het belangrijk dat voldoende aandacht wordt besteed aan deze kwaliteitscriteria. Dat betekent onder andere dat gegevens op een doelmatige manier verzameld worden: er moet met andere woorden over gewaakt worden dat er op een correcte (of betrouwbare) manier juiste (of valide) gegevens worden verzameld en dat op een zo efficiënt mogelijke manier. Daarnaast is het belangrijk dat de evaluatie ook voldoende billijk of rechtvaardig is. We spreken van een rechtvaardige evaluatie indien er sprake is van objectiviteit, doorzichtigheid en normering. In principe zijn deze kwaliteitseisen van belang in elke fase van het evaluatieproces. Toch zijn er accentverschillen in de verschillende fases. Zo is validiteit bijvoorbeeld voornamelijk aan de orde wanneer evaluatietaken ontwikkeld worden. Bij het beoordelen van evaluatiegegevens is dan weer vooral betrouwbaarheid en objectiviteit van belang. Beide aspecten van een goede evaluatie werden reeds uitvoerig toegelicht in deel 2. De aldaar geformuleerde eisen zijn ook hier onverminderd van toepassing.

In de volgende paragraaf stellen we ons de vraag op welke manier de gegevensverzameling en de beoordeling van evaluatiegegevens vorm moeten krijgen opdat er sprake zou zijn van betrouwbaarheid en objectiviteit.

5.2 Hoe een betrouwbaar waarderingsproces realiseren?

Er zijn verschillende evaluatieaspecten die onze aandacht verdienen wanneer we met de in vorige stappen ontwikkelde taken en instrumenten aan de slag willen gaan. Om de gegevensverzameling en de beoordeling zo betrouwbaar mogelijk te laten verlopen, zijn ons inziens vier aandachtspunten van belang: iedereen moet hetzelfde verstaan onder de criteria; de beoordelaars

moeten geoefend zijn in het gebruik van de criteria; de beoordeling mag niet gebaseerd zijn op een te beperkt aantal gegevens; en er moet een controle-instrument worden voorzien.

1. Verzekeren dat iedereen hetzelfde verstaat onder de criteria

Essentieel om tot een betrouwbaar beoordelingsproces te komen, is het verzekeren van een consensus bij verschillende beoordelaars omtrent de betekenis van criteria en de manier waarop de criteria gebruikt dienen te worden. Een mogelijkheid hiertoe is het ontwerpen van een begeleidend document dat de typische beoordelingsrichtlijnen voor de betreffende evaluatie uiteenzet. Herman et al. (1992) stellen dat een dergelijk document bijvoorbeeld volgende informatie kan bevatten: duidelijk toegelichte beoordelingscriteria; voorbeelden die elk prestatieniveau illustreren; een handige samenvatting van de criteria voor gebruik tijdens de eigenlijke evaluatie; en een formulier om op een systematische manier de evaluatiegegevens te noteren.

2. Train de evaluatoren in het gebruik van criteria

Men mag veronderstellen dat leerkrachten een zekere professionele ontwikkeling moeten doormaken alvorens een betrouwbaar waarderingsproces gegarandeerd kan worden. Het beoordelen van complexe vaardigheden is -omdat er in vele gevallen geen sprake is van één enkel juist antwoord- immers niet eenvoudig. Daarom is het belangrijk dat voldoende aandacht geschonken wordt aan het trainen van beoordelaars in het gebruik van evaluatieprocedures- en criteria (Dochy, 1999b). Te denken valt aan het leren classificeren, interpreteren en gebruiken van criteria, het waarden van een leerprestatie en het legitimeren van een waardeoordeel. Ook de manier waarop waardeoordelen gerapporteerd worden, moet worden geoefend. Dergelijke trainingen zijn zowel leerrijk wanneer ze zich aanvankelijk -eerder algemeen- richten op het gebruik van criteria als wanneer ze zich toelagen op het gebruik van concrete criteria binnen een bepaalde feitelijke evaluatiecontext.

De training in een goed gebruik van beoordelingsinstrumenten is daarnaast ook van belang omdat een aantal veel voorkomende fouten de objectiviteit van evaluatiebeoordelingen in het gedrang kunnen brengen. Een bekende fout is bijvoorbeeld het zogenaamde 'halo-effect': als een leerling op een door de beoordelaar belangrijk geachte gedragscomponent (zoals mooi handschrift) erg goed scoort, zal de beoordelaar de neiging hebben om de leerling ook op andere gedragscomponenten hoog te scoren (zoals zinsbouw en woordgebruik). Beoordelaars kunnen er dus foutief van uitgaan dat bepaalde gedragscomponenten onderling samenhangen en alzo leerlingen op deze elementen ongeveer dezelfde score geven, zonder verder na te gaan hoe de leerlingen effectief op elk van de afzonderlijke gedragscomponenten scoren. Ook het feit dat beoordelaars soms geneigd zijn om enkel de uiteinden van een schaal te gebruiken, maakt een goede training noodzakelijk (Janssens et al., 2000).

3. Gebruik meerdere observatie- en beoordelingsmomenten

Indien mogelijk wordt de beoordeling van leerlingenprestaties best gebaseerd op gegevens van meerdere evaluatietaken en observatiemomenten. Op die manier kan men de generaliseerbaarheid van de beoordeling verhogen. De invloed van toevalligheden en eventuele meetfouten wordt immers minder groot. Dit geldt des te meer wanneer gewichtige beslissingen worden genomen, zoals het al dan niet mogen overgaan naar een volgend schooljaar.

4. Voorzie een controle-instrument

Om zicht te krijgen op de betrouwbaarheid van het evaluatieproces kan men een controlesysteem ontwikkelen dat de consistentie van beoordelingen bewaakt. Hierbij dient de aandacht zowel te worden gevestigd op de consistentie over de tijd heen als op de consistentie over verschillende beoordelaars. Want nogmaals: evaluatoren moeten eenzelfde prestatie op verschillende evaluatiemomenten en -contexten gelijk beoordelen.

Er zijn verschillende manieren om mogelijke controlemechanismen in het evaluatieproces in te bouwen. Te denken valt bijvoorbeeld aan de volgende:

- ✓ Ontwikkel voor elke score die kan worden gegeven een exemplarisch voorbeeld. Dit betekent dat men voor elke score een voorbeeld maakt of deze uit reeds beschikbare evaluaties verzamelt. Deze controlevoorbeelden maken het mogelijk om de consistentie van de beoordeling over beoordelaars en over de tijd na te gaan en eventueel bij te sturen.
- ✓ Om de consistentie over de tijd te bewaken kan men tevens reeds vroeger beoordeelde evaluatietaken (van vorig schooljaar bijvoorbeeld) samen met de nieuwe te evalueren taken opnieuw scoren. Op die manier kan men bepaalde evoluties op het spoor komen.

Ter afsluiting rapporteren we een checklist die het mogelijk maakt om na te gaan of een bepaald waarderingsproces al dan niet voldoet aan de verschillende eisen.

Figuur 3.22: De betrouwbaarheid van een beoordelingsproces beoordelen

Beschikt u over...

- duidelijke, concrete criteria?
- een gedocumenteerde en uitgeteste beoordelingsleidraad?
- een geschikte en betrouwbare gegevensverzameling?
- becommentarieerde voorbeelden voor alle mogelijke beoordelingen?
- voldoende oefening en feedback voor de beoordelaars?
- garanties voor de consistentie van beoordelingen over beoordelaars?
- periodieke controle van de betrouwbaarheid?
- de mogelijkheid tot bijkomende training indien nodig?
- voldoende evaluatiegegevens (gebaseerd op verschillende observatiemomenten)?

6 Stap 6: Communicatie en gebruik van de resultaten

De laatste fase van het evaluatieproces bestaat uit het meedelen van de evaluatieresultaten aan de verschillende betrokkenen. Ook in deze laatste stap moeten nog een aantal keuzes worden gemaakt. Hierbij staan twee vragen centraal: (1) waarvoor zullen de evaluatiegegevens aangewend worden en (2) op welke manier worden de evaluatieresultaten best naar de betreffende doelgroepen gecommuniceerd.

De vraag waartoe de evaluatiegegevens gebruikt zullen worden, is in principe reeds in de eerste stap van het evaluatieproces beantwoord. Het antwoord was op dat moment reeds bepalend voor het verdere verloop van het evaluatieproces. In deze fase van de evaluatie komt het er dus op neer de evaluatiegegevens aan te wenden voor de doelen die in de eerste stap vooropgesteld werden. De verschillende doelen van evaluatie werden in deel 1 reeds beschreven. Een aantal belangrijke aspecten bij het gebruik van de evaluatiegegevens verdienen evenwel nog even onze aandacht. In vraagvorm geformuleerd komt het hierop neer: op welke manier worden de verzamelde evaluatiegegevens best aan de betreffende doelgroepen bekend gemaakt? Om deze vraag bij elke evaluatie te kunnen beantwoorden moet worden stilgestaan bij volgende punten: (1) welke scoringsstrategie is aangewezen en (2) wat is de meest geschikte manier om de evaluatiegegevens te communiceren.

6.1 Scoringsstrategieën

Een van de fundamentele beslissingen die moet genomen worden, betreft de manier waarop de leerlingprestaties zullen worden beoordeeld. Drie verschillende strategieën worden in de huidige evaluatiepraktijk onderscheiden. Deze zijn relatieve scoring, criteriumgerichte scoring en ipsatieve scoring.

1. Relatieve scoring

Bij relatieve scoring (normgerefeerde of normatieve scoringsstrategie) worden de prestaties van een individuele leerling vergeleken met de prestaties van anderen. Er wordt als het ware een rangschikking gemaakt van de prestaties van alle leerlingen (Bond, 1996). Vervolgens wordt aan de leerlingen met de beste prestatie een hoge score toegekend en aan de leerlingen met de minst goede prestatie een lage score. Hierbij wordt geen rekening gehouden met de vraag of iedereen misschien goed of slecht presteerde. Indien alle leerlingen goed presteerden, zal diegene met de minst goede prestatie toch een lage score krijgen. Niet iedereen kan dus een hoge of een lage score hebben. Voorbeelden van deze scoringsstrategie zijn onder andere het weergeven van percentielen en rangschikkingen van leerlingen (in de vorm van de zogenaamde plaats in de klas). De score van leerlingen is dus gebaseerd op de onderlinge vergelijking van leerlingen en niet op de eigenlijke kwaliteit van hun prestatie.

De normatieve scoringsstrategie gaat uit van de veronderstelling dat de evaluatieprestaties van alle leerlingen een normale verdeling kennen (dat wil zeggen dat de meerderheid van de leerlingen rond het gemiddelde presteert en dat een minderheid zeer goed of ondermaats presteert). Deze assumptie is om volgende twee redenen problematisch (Mabry, 1999). *Ten eerste*, in het onderwijs (c.q. met de eindtermen) gaat men ervan uit dat een meerderheid van de leerlingen een vooropgestelde prestatie met succes moet kunnen neerzetten en dat dus de meeste (en de rationale van de eindtermen volgend zelfs alle) leerlingen een hoge score moeten behalen. Of anders geformuleerd: indien leerlingen gerangschikt worden op basis van een normale verdeling, dan impliceert dit dat steeds de helft van de leerlingen onder de gemiddelde verwachting scoort. Deze werkwijze leidt dus steeds tot ontgoocheling en ontmoediging bij een belangrijk deel van de leerlingen. *Ten tweede* moet worden benadrukt dat normatieve scoring leerlingen als het ware concurrenten maakt van mekaar. Wanneer één leerling een betere score krijgt, zullen anderen immers automatisch een slechtere score krijgen. Deze scoringsstrategie is dan ook contraproductief, in die zin dat leerlingen worden ontmoedigd om van mekaar te leren en om in groep samen te werken.

Ook al is het interessant om te weten hoe leerlingen scoren in verhouding tot hun leeftijds- of klasgenoten, een rangschikking geeft geen informatie over het prestatieniveau dat leerlingen bereiken en over wat ze geleerd hebben. En ook al leest men soms dat competitie kan stimuleren om harder zijn best te doen, toch is dit enkel het geval voor de besten van de klas terwijl de zwakkeren ontmoedigd worden. Bovendien zou competitie en vergelijking de intrinsieke motivatie kunnen doen dalen (Janssens et al., 2000; De Neve & Janssen, 1992). Deze scoringsstrategie past dus niet binnen een assessmentcultuur.

2. Criteriumgerichte scoring

De prestaties van leerlingen worden bij criteriumgerichte scoring (ook absolute of standaardgebaseerde scoring) afgewogen tegen vooraf bepaalde criteria en standaarden. Er wordt bij de beoordeling uitgegaan van de vooropgestelde doelstellingen. Hoge scores worden toegekend aan leerlingprestaties die tegemoet komen aan of beter zijn dan de vooropgestelde standaarden. Lage scores worden toegekend aan prestaties die -afgezet tegen de criteria- onvoldoende kwaliteitsvol zijn. Een mogelijk criterium is bijvoorbeeld maximum vier fouten schrijven in een dictee. In dat geval is het mogelijk dat iedereen slaagt, maar ook dat niemand slaagt.

In principe zet deze scoringsstrategie het evaluatieproces verder in de geest waarop het tot op heden besproken is. Aan de hand van de criteria weten leerlingen immers vooraf welke prestatie zal resulteren in een hoge of lage score. Voorbeelden van deze scoringsstrategie zijn bijvoorbeeld scores gebaseerd op rubrieken en evaluatielijsten.

3. Ipsatieve scoring

Bij de ipsatieve (of leerlinggerichte) scoringsstrategie worden de prestaties van leerlingen niet vergeleken met een extern referentiepunt (zoals in de vorm van criteria en standaarden in de

criteriumgeriefereerde scoring), maar eerder met een intern individueel vergelijkingspunt. Dit punt is gebaseerd op persoonlijke sterktes en zwaktes van de leerling, op diens reeds verworven vaardigheden en op diens vooruitgang, leerkansen, interesses, doelstellingen, achtergrond en andere factoren die van invloed kunnen zijn op de betreffende prestatie. Een zelfde prestatieniveau kan dus bij verschillende leerlingen tot een verschillende score leiden.

Bij ipsatieve evaluatie is het dus zeker niet de bedoeling leerlingen met mekaar te vergelijken. Men beoogt daarentegen een volledig inzicht te verwerven van elke afzonderlijke leerling en een beoordeling van diens unieke verwezenlijkingen (Mabry, 1999). Criteria worden hiertoe niet gezocht in het curriculum. Het is immers perfect mogelijk dat een leerling meer weet van een bepaald onderwerp dan het curriculum vooropstelt. Of aan het andere uiterste, indien een leerling een kleine vooruitgang maakt van een zeer slechte prestatie naar een slechte prestatie zullen normatieve en criteriumgerichte scoringsstrategieën deze moeilijk aan het licht kunnen brengen. Ipsatieve scoring zal bij de betreffende leerling wel een verbetering constateren, met alle belangrijke positieve gevolgen van dien. Deze scoringsstrategie kan dus best aangewend worden indien een verregaande differentiëring en een intensieve begeleiding van leerlingen aangewezen is.

6.2 De communicatie van evaluatieresultaten

Feedback en de rapportage van evaluatiegegevens kunnen op verscheidene manieren plaats vinden. Leerkrachten, leerlingen en ouders zijn echter zo vertrouwd met het geven van puntscores dat ze moeilijkheden ondervinden om zich voor deze praktijk alternatieven voor te stellen (Mabry, 1999). Meer en meer gaan er echter stemmen op om de communicatie over evaluatieresultaten meer leerlinggeoriënteerd te maken. In een eerste paragraaf geven we een overzicht van een aantal werkwijzen om evaluatiegegevens en feedback te communiceren. Daarna staan we stil bij de keuze van een geschikte werkwijze.

6.2.1 Hoe evaluatiegegevens rapporteren?

Het terugkoppelen van evaluatiegegevens gebeurt in de huidige Vlaamse onderwijspraktijk veelal in de vorm van een score op een 10-puntenschaal of in de vorm van percentages. Op een examenrapport wordt er in het verlengde van deze praktijk melding gemaakt van de punten voor elke toets, de percentages, het klasgemiddelde, een totaal van alle vakken, een rangorde van de leerlingen, maar ook van begeleidingsvoorstellen en een advies van de klassenraad (Verhoeven et al., 2000). Naast het louter toekennen van punten zijn er dus aanzetten om een aantal alternatieve werkwijzen te gebruiken. Hieronder besteden wij aandacht aan deze werkwijzen. Respectievelijk komen aan bod: de verschillende soorten schalen, evaluatielijsten en rubrieken, en narratieve rapporten.

1. De verschillende soorten schalen

Voor sommige onderwijsdoelen kan het volstaan dat wordt aangegeven of leerlingen bepaalde taken al dan niet uitgevoerd hebben of dat de prestatie op de taak al dan niet volstaat. De communicatie van evaluatiegegevens zou hier evenwel niet toe beperkt mogen worden. Voor leerlingen is het zinvol dat er een uitgebreider waardeoordeel wordt uitgesproken over hun prestatie. Hiertoe werden in de vierde stap verschillende soorten beoordelingsschalen gepresenteerd. We hebben daar aangetoond dat het loont om op zoek te gaan naar alternatieven voor het traditionele puntengebruik.

2. Evaluatielijsten en rubrieken

Een werkwijze die het mogelijk maakt om gedetailleerder over de prestaties van leerlingen te rapporteren, bestaat uit het gebruik van de beoordelingsinstrumenten die in de vroegere stappen werden ontwikkeld, namelijk de evaluatielijsten en rubrieken. Ook aan deze elementen hoeven we niet veel aandacht meer te besteden. We onthouden voornamelijk dat het terugkoppelen van zeer gedetailleerde evaluatiegegevens het gebruik van analytische rubrieken en evaluatielijsten vereist. Meer algemene evaluatiegegevens kunnen gerapporteerd worden aan de hand van een beoordeling op een holistische rubriek.

3. Een narratief rapport

Narratieve (of ook anekdotische) rapporten verschillen van andere rapporteringstechnieken van evaluatiegegevens omwille van het feit dat ze meer uitgebreid en informatief zijn. Voor leerlingen, ouders en leerkrachten is het zeer interessant te beschikken over verslagen die aangeven welke vaardigheden leerlingen al dan niet beheersen. Scores op beoordelingsschalen zijn soms immers onvoldoende om leerlingen en hun ouders een goed beeld te geven van wat de leerling reeds bereikt heeft (Wiggins, 1994). Stel dat uw dochter een B krijgt voor spelling, dan leert deze beoordeling dat zij over het algemeen een goede speller is. Maar wat als uw dochter systematisch fouten maakt tegen het al dan niet schrijven van de 's' als verbindingsletter. Met een B blijft u verstoken van deze informatie. Een narratief rapport kan dergelijke elementen echter op een eenvoudige manier aan het licht brengen.

Figuur 3.23: Een voorbeeld van een narratief rapport

Volgend voorbeeld past binnen het vak Nederlands en is geschreven voor Pieter, een leerling uit de eerste graad secundair onderwijs.

Pieter, je hebt de laatste maanden bij een aantal aspecten van Nederlands grote vooruitgang geboekt, voornamelijk wat zinsbouw en spelling betreft. Je mondelinge taalgebruik was reeds zeer goed en is dat nog steeds. Een paar dingen wil ik je over schrijven en spelling meegeven. Het verheugt me dat je steeds meer zelfvertrouwen toont wanneer je schrijft. Je vooruitgang op verschillende punten zal daar allicht niet vreemd aan zijn: het gebruik van leestekens, de goede zinsconstructies en het feit dat je op zoek gaat naar de juiste woorden bijvoorbeeld. Ook het feit dat je er een gewoonte van hebt gemaakt om wat je geschreven hebt na te lezen alvorens het in te dienen, maakt dat het een meer verzorgde indruk geeft. Soms worden nog wel eens letters en woorden van plaats verwisseld. Neem je wel genoeg tijd als je je teksten controleert? Nog één ding: je schrijft regelmatig verschillende losse

korte zinnen na mekaar. Door deze zinnen samen te voegen met hoofd- en bijzinnen kan je de teksten vlotter leesbaar maken. Daar moet je de komende weken zeker aandacht aan besteden.

De troef van narratieve rapporten is aan de hand van geschreven of mondelinge commentaren duidelijk en direct te communiceren met leerlingen over sterke en zwakke elementen, over de verwachte prestaties, over de noodzakelijke verbeteringen etc. Op een positieve en informatieve manier geschreven kunnen narratieve rapporten een verscheidenheid aan elementen aan het licht brengen terwijl de waardigheid van de leerling gegarandeerd wordt. Deze manier om evaluatiegegevens te rapporteren, is dus zeker aangewezen wanneer er sprake is van leer- of gedragsproblemen (Brualdi, 1998). Niet alleen leerlingen hebben echter baat bij narratieve rapporten. Ook leerkrachten kunnen de voordelen ervan genieten. Die bestaan uit een reflectie op het verzorgde onderwijsproces (wat werkt, hoe goed, bij wie en waarom?) en een beter zicht op de prestaties van de leerlingen.

De uitgebreidheid en de structuur van een narratief rapport kunnen sterk verschillen. Het is dus niet zinvol om een vast format te presenteren. Vragen betreffende de keuze of de rapportage al dan niet vakoverschrijdend moet worden georganiseerd en welke aspecten in het rapport benadrukt moeten worden, kunnen best bij elke evaluatie opnieuw beantwoord worden. Dit neemt niet weg dat een aantal richtlijnen kunnen worden vooropgesteld.

- In het voorbeeld in figuur 3.23 worden in het narratieve rapport beoordelingen gekoppeld aan de observaties die de leerkracht deed. De leerkracht beperkt zich dus niet tot de feitelijke beschrijving van gedragingen en feitelijke gebeurtenissen in de klas. Of dit al dan niet aangewezen is, hangt af van het doel van je rapportage. Probeer indien mogelijk de narratieve commentaar neer te schrijven. Op die manier kan later naar de gegevens worden teruggegrepen en kunnen evoluties in het leerproces aan het licht komen. Daartegenover staat dat mondelinge rapportage veel minder tijdsintensief is.
- Ongeacht de vorm van het narratieve verslag, tracht zoveel mogelijk specifieke commentaar te geven. Gerichte en gedetailleerde feedback is immers informatiever dan algemene beoordelingen. Het feit dat een leerling leergierig, trots, luisterbereid, doordacht en zelfstandig is, leert veel meer over die leerling dan enkel te stellen dat hij/zij een goede leerling is. Een goede observatie en het stellen van de juiste vragen zijn dan ook noodzakelijk. Uiteraard dienen narratieve rapporten met de nodige zorg te worden samengesteld. Wanneer ze gereduceerd worden tot feedback zoals 'goed gewerkt' of 'werk aan de winkel' zijn geschreven commentaren immers weinig informatiever dan cijferscores of graden zonder kwalitatieve beoordeling. Leerlingen en hun ouders moeten met de informatie aan de slag kunnen. De evaluatiegegevens moeten daarom inspelen op toekomstige activiteiten en beslissingen (Morris et al., 1987).
- Narratieve rapporten kunnen dagelijks voor elke leerling worden bijgewerkt, maar gezien de tijdsinvestering die dit vereist, gebeurt de rapportage veelal minder frequent (gaande van bijvoorbeeld een- of tweemaal per week tot eenmaal per trimester). Sommige leerkrachten werken met post-its waarop ze onmiddellijk belangrijke observaties

noteren die ze dan achteraf in het dossier van de betreffende leerling plakken. Minder tijdsintensief is dan weer het gebruik van een checklist met leerlingengedragingen die wordt aangevuld met korte vaststellingen.

De resultaten van leerlingen worden momenteel na het afsluiten van een lesperiode samengebracht in een zogenaamd rapport, dat veelal bestaat uit een opsomming van puntenscores voor de verschillende vakken. Vaak is in deze rapporten te zien hoe de leerling heeft gepresteerd vergeleken met leeftijdgenoten. We toonden aan dit maar één mogelijke (en bovendien bekritiseerbare) strategie is om leerlingprestaties te beoordelen. De puntenscores blijken bovendien in vele gevallen niet de meest aangewezen manier om evaluatiegegevens te communiceren. Het komt er dan ook op aan een doordachte keuze te maken tussen de verschillende rapporteringswijzen.

6.2.2 De keuze van een rapporteringswijze

Evaluatiegegevens kunnen op verschillende manieren gecommuniceerd en gebruikt worden. Elke manier beoogt eigen doelen en heeft specifieke voor- en nadelen. De keuze van de meest geschikte communicatiemethode van de verzamelde evaluatiegegevens is afhankelijk van de aard van specifieke kenmerken van de evaluatie. De centrale vraag luidt: wie moet wat weten, en wanneer? Ook hier geldt dus dat er niet één beste methode bestaat om alle evaluatiegegevens te rapporteren. De verscheidenheid aangaande de inhoud van de evaluatie, de doelen van de evaluatie, de evaluatietaken, de beoordelaars en het publiek maken dat verschillende communicatiemethoden moeten worden aangewend. Ook de beschikbare tijd en middelen zijn uiteraard van belang. Het komt er op aan een duidelijk beeld te hebben van welke informatie de gebruikers nodig hebben en waarvoor deze informatie gebruikt zal worden. Bij de meeste evaluaties verdient het aanbeveling een combinatie van bovenstaande mogelijkheden voorop te stellen. De ene werkwijze is immers geschikt om algemene beoordelingen uit te spreken terwijl de ander gepast is bij het rapporteren van feedback. Het maken van een keuze betekent niet dat er steeds móet gekozen worden tussen cijfers, graden, kwalitatieve beoordelingsschalen en narratieve rapporten. Vaak is het aangewezen een combinatie te gebruiken. Graden en beoordelingsschalen zullen narratieve rapporten immers eerder aanvullen dan vervangen. Kritieken op de traditionele cijfer rapporten pleiten er voor om meer uitgebreide informatie over de leerling op te nemen in de rapporten. De zogenaamde eenduidigheid en exactheid van de 10-puntenschaal en de procentuele puntentoekenning werden immers sterk bekritiseerd. Een goed rapport moet dan ook meer omvatten. Het is belangrijk om duidelijke en voldoende informatie te verschaffen waarmee de leerling aan de slag kan. Hiertoe moeten de mogelijkheden van de verschillende rapporteringswijzen gecombineerd worden tot een passend geheel.

7. Opmerkingen bij het voorgestelde evaluatieproces

Met de beschrijving van de verschillende rapporteringswijzen lijkt het alsof we aan het einde zijn gekomen van een afgerond geheel. Er werden evaluatiedoelen vooropgesteld, taken en evaluatieinstrumenten ontwikkeld, gegevens verzameld en beoordeeld en de evaluatiegegevens werden teruggekoppeld naar de verschillende betrokkenen. Op dat punt zouden dus nieuwe leer- en dus ook evaluatiedoelen kunnen worden vooropgesteld. Het evaluatieproces kan dan opnieuw van start gaan. Deze indruk van mooi opeenvolgende cycli vloeit voort uit de chronologische benadering die we hanteerden om het evaluatieproces te beschrijven. Toch is deze indruk als het ware misleidend. Om tot een goed begrip te komen van het voorgestelde evaluatieproces moeten daarom nog twee zaken extra benadrukt worden, namelijk de flexibiliteit en de contextafhankelijkheid van het evaluatieproces.

Het evaluatieproces is niet-lineair en moet dus flexibel worden toegepast. Beslissingen in elke stap van het proces worden immers zowel beïnvloed door de stappen die er aan voorafgaan als door diegene die er op volgen. In die zin is de manier waarop (met de onderscheiden stappen) het evaluatieproces in het vorige werd beschreven enigszins artificieel. Dat blijkt onder meer uit de volgende voorbeelden.

- ✓ Tijdens het ontwerpen van interessante taken is het mogelijk dat men nog interessante doelstellingen op het spoor komt die men oorspronkelijk over het hoofd heeft gezien. Het is dus niet steeds zo dat de eerste stap van het proces volledig is afgerond alvorens men verder gaat.
- ✓ De stappen 'ontwikkelen van evaluatietaken' en 'ontwikkelen van criteria' zijn vaak enkel in theorie volledig van mekaar te onderscheiden.
- ✓ Het ontwikkelen van evaluatiecriteria in stap 4 is dan weer onlosmakelijk verbonden met het garanderen van een betrouwbaar evaluatieproces in stap 5.
- ✓ Argumenten die van belang zijn bij het ontwerpen van een geschikte beoordelingsschaal komen vaak pas aan het licht wanneer men de gegevens naar de betrokkenen gaat terugkoppelen. Men moet dus eigenlijk reeds in stap 4 reflecteren over stap 6.

Met contextafhankelijkheid van het evaluatieproces bedoelen we dat de noodzaak van een strikte toepassing van de verschillende stappen in het proces afhankelijk is van de doelen en de gevolgen van de evaluatie. Het is dus belangrijk te benadrukken dat er geen algemeen geldende of beste manier bestaat om een evaluatie uit te voeren. De noodzaak om alle stappen uit het evaluatieproces even nauwgezet uit te voeren, is dan ook niet altijd even belangrijk. Veel hangt af van de concrete context, de evaluatoren en de objecten van de evaluatie. Zo staan bij het bepalen van de vereiste validiteit en betrouwbaarheid van een evaluatie de doelen en de gevolgen van de betreffende evaluatie centraal. De vraag welke doelen men met een specifieke evaluatie wenst te bereiken, kan worden geconcretiseerd in verschillende deelvragen: wat wil je evalueren? wie wil je evalueren? waarom wil je evalueren (formatief, summatief)? wie gaat de evaluatiegegevens gebruiken? De vragen die dus in de eerste stap van het evaluatieproces aan de orde zijn. De antwoorden op elk van deze vragen kunnen aanzienlijke implicaties hebben voor de inhoud en het verloop van de evaluatie en voor de aandacht die men aan de kwaliteitsvolle uitvoering van de evaluatie dient te besteden. Zo moeten de betrouwbaarheid en validiteit minder

gewaarborgd zijn bij het tussentijds evalueren van een leerling als basis voor gerichte ondersteuning dan bij het maken van eindbeoordelingen op basis waarvan men beslissingen wil treffen over het al dan niet laten overgaan van een leerling naar een volgend schooljaar. Het weze dus duidelijk dat bepaalde beslissingen meer vereisten impliceren dan anderen. Hoe groter de belangen die aan de evaluatie vasthangen, des te belangrijker de kwaliteitsgaranties. Het voorgaande betekent uiteraard niet dat men de kwaliteit van minder gewichtige evaluaties mag verwaarlozen. Ook deze evaluaties (zoals informele voortgangsevaluaties) kunnen immers aanzienlijke schade berokkenen aan leerlingen indien ze onvoldoende verzorgd gebeuren.

Deel



Alternatieve evaluatievormen geïllustreerd

INHOUD - DEEL 4

Alternatieve evaluatievormen geïllustreerd

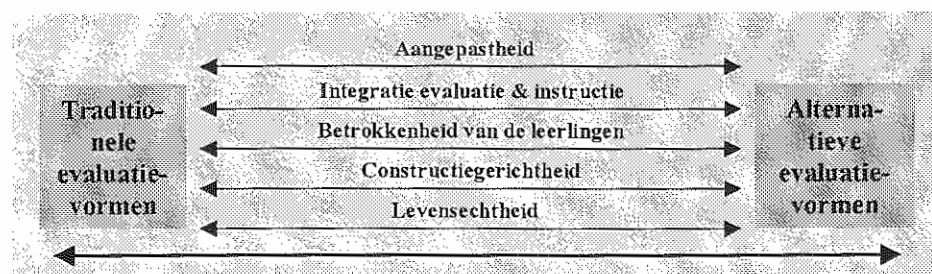
Hoofdstuk 1: Vaardigheidsevaluatie	178
1 Vaardigheidsevaluatie geïllustreerd	178
2 Vaardigheidsevaluatie gedefinieerd	179
3 Kenmerken en doelen van vaardigheidsevaluatie	180
3.1 Kenmerken van vaardigheidsevaluatie	180
3.2 Doelen van vaardigheidsevaluatie	181
4 Vaardigheidstaken	182
4.1 Karakteristieken van vaardigheidstaken	183
4.2 Types vaardigheidstaken	184
4.3 De beoordeling van vaardigheidstaken	185
5 Voorbeelden van vaardigheidstaken	185
5.1 Algemene voorbeelden van vaardigheidstaken	185
5.2 Voorbeelden van vaardigheidstaken in onze casussen	186
Hoofdstuk 2: Authentieke evaluatie	193
1 Authentieke evaluatie geïllustreerd	193
2 Authentieke evaluatie gedefinieerd	194
3 Kenmerken en doelen van authentieke evaluatie	195
3.1 Kenmerken van authentieke evaluatie	195
3.2 Doelen van authentieke evaluatie	195
4 Authentieke evaluatietaken	196
4.1 Hoe vaardigheidstaken authentiek maken?	197
4.2 Inspiratiebronnen voor authentieke taken	198
5 Voorbeelden van authentieke taken	199
5.1 Algemene voorbeelden van authentieke taken	199
5.2 Voorbeelden van authentieke taken in onze casussen	199

Hoofdstuk 3: De betrokkenheid van leerlingen	207
1 Verschillende vormen van leerlingenbetrokkenheid	208
1.1 De betrokkenheid van leerlingen geïllustreerd	209
1.2 Vormen van leerlingenbetrokkenheid gedefinieerd	210
1.3 Kenmerken van de verschillende vormen van leerlingenbetrokkenheid	211
2 Waarom leerlingbetrokkenheid?	213
2.1 Een pleidooi voor leerlingenbetrokkenheid	213
2.2 Toch ook enige terughoudendheid	215
3 Het verloop van leerlingenbetrokkenheid	217
4 Evaluatieinstrumenten	224
4.1 Evaluatielijsten	224
4.2 Attitudeschalen	226
4.3 Logboeken	228
4.4 Mondelinge interviews	232
5 Voorbeelden van leerlingenbetrokkenheid	233
5.1 Algemene voorbeelden van leerlingenbetrokkenheid	234
5.2 Voorbeelden van leerlingenbetrokkenheid in onze casussen	235
Hoofdstuk 4: Het gebruik van portfolio's	248
1 Portfolio geïllustreerd	248
2 Portfolio gedefinieerd	250
3 Kenmerken en doelen van een portfolio assessment	251
3.1 Kenmerken van een portfolio assessment	251
3.2 Doelen van een portfolio assessment	253
4 Waarom portfolio's?	254
5 Het evaluatieproces bij portfolio's	255
5.1 Organisatie en planning	255
5.2 Verzameling – Samenstellen van de portfolio	257
5.3 Reflectie	261
6 Verschillende types portfolio	262
7 De verschillende componenten van een portfolio	263
7.1 Inhoudelijke componenten	263
7.2 Organisatorische componenten	264
8 Het beoordelen van een portfolio	266
9 Voorbeelden van portfoliogebruik	268

Inleidend

In de vorige delen werd duidelijk dat de term 'alternatieve evaluatie' een verzamelnaam is. Verschillende evaluatiemethoden verdienen de benaming alternatieve evaluatie. In deel 2 stelden we een aantal kenmerken voorop van alternatieve evaluatievormen. Hierbij werd benadrukt dat er naargelang het kenmerk of de combinatie van kenmerken die in een bepaalde gerealiseerd worden, sprake is van verschillende alternatieve evaluatiemethoden. Dit betekent dat we een evaluatiemethode die deze verschillende kenmerken weet te combineren als het ware kunnen bestempelen als een volmaakte vorm van alternatieve evaluatie. In dat geval zou de evaluatie immers tegemoet komen aan al de geformuleerde kenmerken. De onderstaande figuur brengt de verschillende kenmerken die maken dat er sprake is van een vorm van alternatieve evaluatie in herinnering.

Figuur 4.1: Continua richting alternatieve evaluatievormen



In het geval bij een bepaalde evaluatie deze verschillende kenmerken (eerder) niet van toepassing zijn, dan zal deze evaluatie aanleunen bij de traditionele evaluatievormen. We willen hierbij benadrukken dat dit niet impliceert dat die evaluatievorm per definitie te verwerpen is. In deel 2 beschreven we immers de kenmerken van een goede evaluatie in het algemeen. Een concreet evaluatieinstrument kan perfect aan deze eisen voldoen -en dus een goede evaluatie zijn- zonder kenmerken van alternatieve evaluatievormen in zich te dragen. Wanneer wel wordt tegemoet gekomen aan een of meerdere van bovenstaande kenmerken spreken we van een alternatieve evaluatiemethode.

Elk van de kenmerken kan worden opgevat als een continuüm. Of een evaluatiemethode al dan niet alternatief is, is dus geen kwestie van alles of niets. Dit maakt 'alternatief' tot een omvattend en gradueel begrip. **Omvattend** omdat ook evaluatievormen die niet aan al de vooropgestelde eisen voldoen maar daarentegen enkel één kenmerk benadrukken als 'alternatief' bestempeld kunnen worden (te denken valt aan evaluatievormen zoals vaardigheidsevaluatie en zelfevaluatie). Niet elk kenmerk moet dus in dezelfde mate gerealiseerd worden om te kunnen spreken van een alternatieve evaluatievorm. Een evaluatiemethode die zich expliciet op één kenmerk van alternatieve evaluatievormen toelegt, laat de andere kenmerken vaak (gedeeltelijk) voor wat ze zijn. Een evaluatiebenadering die gebruik maakt van levensechte, authentieke taken

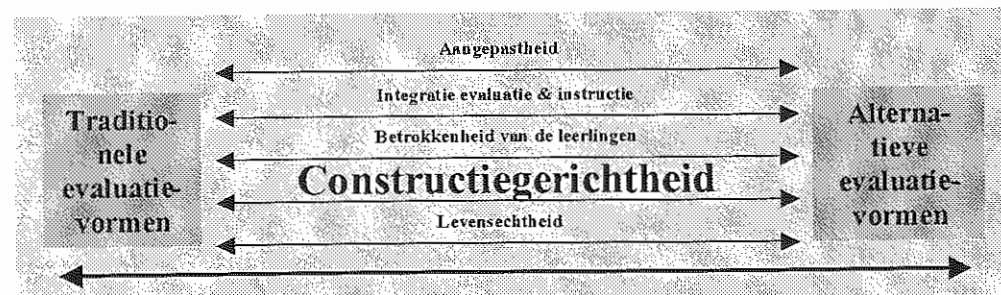
hoeft bijvoorbeeld niet per se eveneens de betrokkenheid van leerlingen aan te halen. **Gradueel** omdat het toekennen van het begrip 'alternatief' afhankelijk is van het feit in welke mate een bepaalde evaluatievorm tegemoet komt aan een aantal kenmerken die zelf als een continuüm kunnen worden voorgesteld.

In het vierde deel van dit handboek beschrijven we een aantal evaluatievormen die de benaming 'alternatieve evaluatievorm' verdienen. De verschillende hoofdstukken in dit deel presenteren telkens een evaluatiemethode die één of een aantal kenmerken van alternatieve evaluatievormen benadrukt. Elk van die evaluatiemethoden gaat dus voldoende ver om als alternatieve evaluatievorm te worden bestempeld. Bij elke evaluatiemethode geven we aan welke kenmerken maken dat deze benaming terecht is. Achtereenvolgens komen vaardigheidsevaluatie, authentieke evaluatie, zelf- en peerevaluatie, collaboratieve evaluatie en het gebruik van portfolio's aan bod.

Hoofdstuk 1: Vaardigheidsevaluatie

Het feit dat we vaardigheidsevaluatie als eerste alternatieve evaluatievorm voorstellen, is geen toeval. Vaak wordt de constructiegerichtheid van een evaluatie als essentieel en noodzakelijk kenmerk naar voren geschoven om te kunnen spreken van een alternatieve evaluatievorm. Vaardigheidsevaluatie en alternatieve evaluatie worden in de evaluatieliteratuur dan ook vaak als synoniemen gebruikt. Zoals reeds vermeld is onze benadering evenwel anders: vaardigheidsevaluatie is een vorm van alternatieve evaluatie maar valt er niet mee samen. Dit neemt niet weg dat constructiegerichtheid inderdaad het meest typerende kenmerk is van alternatieve evaluatievormen.

Figuur 4.2: Vaardigheidsevaluatie als alternatieve evaluatievorm



1 Vaardigheidsevaluatie geïllustreerd

De concepten van vaardigheidsevaluatie zijn niet nieuw. Sommige leerkrachten ontwikkelen al jaren taken die perfect kunnen worden ingepast in deze evaluatievorm. Alvorens vaardigheidsevaluatie uitvoeriger te bespreken, willen we met een aantal voorbeelden van vaardigheidstaken (of ook prestatietaken) deze evaluatiemethode illustreren. Centraal in vaardigheidsevaluatie staan immers de vaardigheidstaken die gehanteerd worden.

Figuur 4.3: Voorbeelden van een vaardigheidstaak

Schrijven naar een pennenvriend (uit het onderwijs Nederlands—schrijfvaardigheid)

Lees deze brief van een pennenvriend en beantwoord hem. Vermeld in de brief het volgende: naam, leeftijd, familie, woonplaats, sport, lievelingsmuziek, ... (Popham, 1999).

Hoi,

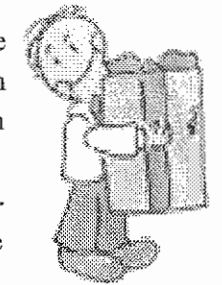
Mijn naam is Peter en ik ben jouw nieuwe pennenvriend. Ik ben dertien jaar oud en heb één zus (Mieke). Mijn moeder en vader zijn allebei onderwijzers. Ik heb zelfs bij mijn vader in de klas gezeten vroeger. We wonen in Ginderbuiten, een gehucht van Mol. Ik vind dat hier weinig te beleven valt en ik verveel me hier soms. Ik zou liever in de stad wonen.

Ik heb verschillende hobby's. Ik verzamel bierkaartjes (ik heb er al 850 verschillende!). Ik ben niet zo gek van sport maar kan wel goed tafeltennissen. Mijn vrienden kunnen mij niet verslagen. Mijn favoriete muziek is rock en mijn favoriete groep is Gorki. Hoe zit het met jou? Schrijf me eens over jouw hobby's en familie.

Berekenen van oppervlakte (uit het rekenonderwijs)

Sandra moet een groot cadeau voor haar zusters verjaardag inpakken. De doos is 60 cm lang, 50 cm breed en 40 cm hoog. In de winkel besluit ze een rol inpakpapier te kopen (met de afbeelding van de favoriete zangeres van haar zus). Op de rol zit 3 vierkante meter papier.

Heeft Sandra volgens jou een verstandige keuze gemaakt door deze rol inpakpapier te kopen? Verklaar hieronder jouw antwoord. Geef daarbij alle wiskundige bewerkingen weer die jouw antwoord ondersteunen.



De zorgsituatie bij de oudere zorgvrager (uit de studierichting Personenzorg)

Stel je eens voor dat je als bejaarde voor de eerste maal in het rusthuis hulp behoeft bij het ochtendtoilet. Wat zou jij dan van de verzorgenden verlangen en verwachten om dit voor jou zo aangenaam mogelijk te laten verlopen? Schrijf hieronder een aantal richtlijnen voor de verzorgenden zodat ze zoveel mogelijk rekening kunnen houden met jouw eigen behoeftes en verlangens (te denken valt onder andere aan vroegere gewoontes, privacy en veiligheid, zelfredzaamheid en dergelijke meer).

2 Vaardigheidsevaluatie gedefinieerd

De definities van vaardigheidsevaluatie zijn talrijk. In het algemeen worden in de verschillende definities evenwel volgende elementen opgenomen (Callison, 1998; Thio, 1998; Gredler, 1999; Mabry, 1999; Campbell et al., 2000).

Vaardigheidsevaluatie is een evaluatievorm waarbij van de leerlingen verwacht wordt dat ze een antwoord construeren (in tegenstelling tot het kiezen van een antwoord uit een reeks mogelijkheden). Typisch aan vaardigheidsevaluatie is dus dat de leerlingen gevraagd worden om (1) bepaalde vaardigheden te demonstreren door iets te doen dat door de evaluatoren wordt geobserveerd op het moment dat het zich voordoet (een presentatie houden, een stopcontact aansluiten, een spel voor bejaarden begeleiden) of om (2) een taak te vervullen die resulteert in het creëren van een product (dit kan variëren van een kort geschreven antwoord tot een heuse tentoonstelling). Er wordt geëvalueerd of leerlingen de geleerde vaardigheden en competenties succesvol kunnen aanwenden.

Leerlingen moeten aan de slag. Ze worden geobserveerd op het moment dat ze bepaalde vaardigheden demonstreren of ze construeren producten die indicatief zijn voor het al dan niet beheersen van bepaalde kennis en vaardigheden. Eerder dan elke vaardigheid afzonderlijk te evalueren worden taken en opdrachten ontwikkeld die de leerlingen stimuleren de vooropgestelde vaardigheden tezamen te demonstreren. Een voorbeeld uit het veld van de muziek kan dit illustreren. Een musicus wordt immers niet zozeer beoordeeld op basis van hoe goed hij bepaalde toonladders kan spelen, maar eerder op basis van hoe goed hij alle noodzakelijke vaardigheden

samenbrengt in één muziekstuk of concert. Ook de leerling die een brief schrijft, moet kennis en vaardigheden betreffende doelgerichtheid, publiekgerichtheid, tekstconventies, woordgebruik, spelling en stijl kunnen integreren. Het feit dat verschillende vaardigheden in hun samenhang worden geëvalueerd, hoeft echter niet te betekenen dat men geen aandacht kan hebben voor afzonderlijke deelvaardigheden. Dat en andere elementen uit de definitie zullen duidelijker worden bij de verdere bespreking van vaardigheidsevaluatie.

3 Kenmerken en doelen van vaardigheidsevaluatie

3.1 Kenmerken van vaardigheidsevaluatie

Het essentiële kenmerk van vaardigheidsevaluatie kan kort en krachtig als volgt omschreven worden: leerlingen moeten zelf antwoorden en/of producten construeren. Vaardigheidsevaluatie wordt dan ook voornamelijk gekenmerkt door de specifieke evaluatietaken (prestatietaken of vaardigheidstaken genoemd) die gehanteerd worden. Deze evalueren niet zozeer de reproductie van kennis, maar richten zich veeleer op het evalueren van vaardigheden en attitudes. We zouden vaardigheidsevaluatie dan ook kunnen kenmerken als een systematische poging om de toepassingsvaardigheden van de lerende te evalueren. De vraag is of de lerende in staat is om reeds geziene kennis te gebruiken bij het oplossen van nieuwe problemen. Met andere woorden: kennisconstructie is het uitgangspunt, niet kennisreproductie.

Even stilstaan... 4.1: Constructiegerichtheid in Vlaamse scholen

Vlaamse leerlingen construeren zonder te construeren. Kan u na het lezen van onderstaande toelichting akkoord gaan met deze enigszins vreemde stelling? Waarom wel/niet?

Vaardigheidsevaluatie is wellicht de alternatieve evaluatievorm die reeds het meest is ingeburgerd in het Vlaamse onderwijs. Althans dit lijkt op het eerste zicht zo. Enerzijds is het correct te stellen dat het gebruik van evaluatieopdrachten waarbij leerlingen louter antwoorden moeten selecteren in het Vlaamse onderwijs sowieso nooit echt veel aanhang heeft genoten. Er wordt eerder een beroep gedaan op open vragen die een kort antwoord behoeven. Leerlingen moeten dus in principe reeds langer antwoorden construeren. Anderzijds is het zo dat leerlingen het antwoord op de evaluatievragen veelal kunnen bekomen door reproductie. Leerlingen construeren met andere woorden een antwoord dat vaak berust op reproductie. Het toepassingsgerichte aspect van vaardigheidsevaluatie is dus nog niet in de constructie opgenomen. Vandaar: Vlaamse leerlingen construeren zonder te construeren.

Bij het presenteren van de kenmerken van de verschillende alternatieve evaluatievormen moet worden opgemerkt dat in de evaluatieliteratuur vaak ook kenmerken van de assessmentcultuur als typerend voor een specifieke evaluatievorm worden vooropgesteld. Dat betekent in het geval van vaardigheidsevaluatie dat vaak niet alleen het gebruik van constructiegerichte evaluatietaken als kenmerkend voor het gebruik van vaardigheidsevaluatie vooropgesteld wordt. Het con-

structiegerichte wordt in de praktijk immers vaak gecombineerd met allerlei kenmerken van de assessmentcultuur en van andere alternatieve evaluatievormen. Ook bij vaardigheidsevaluatie is het bijvoorbeeld mogelijk om de evaluatie en de instructie te integreren en om leerlingen zoveel mogelijk bij het evaluatieproces te betrekken. Vermeldenswaardig is verder dat voor de eigenlijke evaluatie van vaardigheidstaken meestal rubrieken worden ontworpen (zoals we in stap 3 van het evaluatieproces bespraken) en dat evaluatiecriteria vooraf bekend zijn. De leerlingen weten in dat geval dus goed wat er van hen verwacht zal worden en hebben de kans zich daar op voor te bereiden. Kennis van de criteria die gebruikt zullen worden om de prestaties van leerlingen te evalueren, maakt dat leerlingen zich kunnen oefenen in die vaardigheden en domeinen die ze nog niet voldoende beheersten. Indien mogelijk wordt bij vaardigheidsevaluatie eveneens geprobeerd om de taken zoveel mogelijk levensecht te maken. Vaardigheidsevaluatie en authentieke evaluatie worden dus vaak geïntegreerd in dezelfde evaluatie. Toch maken wij duidelijk het onderscheid tussen beide evaluatievormen (zie hoofdstuk 2).

De kenmerken die de eigenheid uitmaken van vaardigheidsevaluatie worden in de evaluatieliteratuur dus vaak in één adem genoemd met de kenmerken van goede evaluatie in het algemeen en van een assessmentcultuur en alternatieve evaluatievormen in het bijzonder. Deze evaluatiekenmerken zijn in onze benadering echter niet strikt noodzakelijk om van vaardigheidsevaluatie te kunnen spreken. Dat maakt deze kenmerken daarom niet minder nastrevenswaardig.

3.2 Doelen van vaardigheidsevaluatie

Het voornaamste doel van vaardigheidsevaluatie is het verzamelen van relevante evaluatiegegevens. Daarnaast blijkt dat vaardigheidstaken op zich een motiverend effect kunnen hebben op de leerlingen.

1. De juiste evaluatiegegevens verzamelen

Zoals bij elke evaluatie is het voornaamste doel van vaardigheidsevaluatie het verzamelen van gegevens die het mogelijk maken om de prestaties en de vooruitgang van leerlingen ten aanzien van vooropgestelde doelen te beoordelen. Vaardigheidsevaluatie heeft als specifieke doel te evalueren in welke mate leerlingen reeds geleerde concepten en vaardigheden kunnen toepassen in nieuwe situaties. Er wordt nagegaan in welke mate leerlingen zelfstandig problemen kunnen oplossen en kritisch kunnen denken (Callison, 1998). Door leerlingen zelf een antwoord te laten construeren en door hen hun oplossingsproces te laten weergeven, kan men als leerkracht de mogelijkheden van leerlingen evalueren zonder zich enkel te moeten baseren op het al dan niet juist (gekozen) zijn van het antwoord (Chapman, 1990). In dat opzicht maakt vaardigheidsevaluatie het mogelijk een veelheid aan relevante evaluatiegegevens te verzamelen over wat leerlingen weten en kunnen. Hierbij wordt zowel aandacht besteed aan het proces als aan het product van de leerlingprestatie (Kerka, 1995). Bovendien worden leerlingen benaderd als individuen die elk een eigen leertraject bewandelen.

Uit de kenmerken van vaardigheidsevaluatie blijkt dat de manier waarop gegevens verzameld worden duidelijk verschilt van de traditionele toetswijze. Een belangrijk doel van vaardigheidsevaluatie is het vermijden van de negatieve impact van de traditionele toetsen. Bovendien wil men tegemoet komen aan de invloed van recente leertheorieën. De voornaamste aanleidingen tot deze alternatieve evaluatiebenadering zijn de veranderde onderwijsdoelen en een ontevredenheid over de traditionele evaluaties die doorgaans enkel basiskennis en eenvoudige vaardigheden evalueren.

2. Motiveren van leerlingen

Naast de verzameling van evaluatiegegevens is er nog een ander doel dat men met het gebruik van vaardigheidstaken nastreeft, namelijk het motiveren van leerlingen. De aard van de taken maakt immers dat leerlingen uitgedaagd worden en zich aangesproken voelen door wat ze leren. Op die manier ontwikkelen ze een positievere houding tegenover de school en tegenover leren in het algemeen.

4 Vaardigheidstaken

Centraal in het gebruik van vaardigheidsevaluatie staan de vaardigheidstaken. Leerlingen moeten in plaats van bijvoorbeeld een antwoord te selecteren uit een drietal mogelijkheden zelf hun eigen antwoord construeren. Vaardigheidstaken situeren zich dus verder op het continuüm van de constructiegerichtheid. Volgende denkoefening staat hierbij stil.

Even stilstaan... 4.2: Een continuüm van constructiegerichtheid

Stel dat je als leerkracht de leerlingen hebt willen bijbrengen hoe ze op een succesvolle manier in groep problemen kunnen oplossen. Op een gegeven moment wil je evalueren in welke mate de leerlingen deze doelstelling (reeds) bereikt hebben. Hiertoe kan je onder andere de volgende taken aanwenden (Popham, 1999). Lees ze even door en rangschik ze daarna op het continuüm van weinig naar sterk constructiegericht.

1. De leerlingen beantwoorden een aantal meerkeuzevragen over welke de volgende stap is die genomen moet worden wanneer men in groep problemen wil oplossen.
2. De leerlingen werken in kleine groepen om een nieuw probleem op te lossen. De leerkracht observeert en beoordeelt de prestaties van de verschillende leerlingen.
3. De leerlingen beantwoorden 'waar of vals'-vragen over de beste procedures om in groep een probleem op te lossen.
4. De leerlingen krijgen een nieuw probleem voorgeschoteld. Daarna moeten ze uitvoerig uitschrijven op welke manier een groep tewerk zou moeten gaan om het probleem op te lossen.
5. De leerlingen krijgen een aantal vragen voorgeschoteld over werkwijzen om een probleem gezamenlijk aan te pakken. Op deze vragen moeten ze korte antwoorden geven.

Door het gebruik van vaardigheidstaken wil men leerlingen laten demonstreren dat ze bepaalde taken kunnen uitvoeren. Wat deze taken betreft, verdienen zowel de karakteristieken als de verschillende soorten onze bijkomende aandacht.

4.1 Karakteristieken van vaardigheidstaken

We kunnen een vaardigheidstaak als volgt omschrijven: een taak, probleem, of vraag die van de leerlingen verwacht dat ze antwoorden construeren (in plaats van selecteren). Het betreft bijvoorbeeld taken die van de leerlingen verwachten dat ze strategieën herzien, gegevens organiseren, patronen identificeren, modellen formuleren en generaliseren, partiële en tentatieve oplossingen evalueren, en hun antwoorden rechtvaardigen (Mabry, 1999, p. 18). Ervaringen met vaardigheidstaken hebben geresulteerd in een overzicht van criteria waaraan vaardigheidstaken dienen te voldoen. Op basis van deze ervaringen kunnen we stellen dat een effectieve taak volgende karakteristieken moet hebben (WCER, 1993).

Figuur 4.4: Een aantal karakteristieken van vaardigheidstaken

- **Uitlokken van probleemoplossend gedrag, redeneren en communicatie**
De taak is zo opgesteld dat ze niet simpelweg kan worden opgelost door het zich herinneren van een antwoord of het toepassen van een algoritme.
Het denkproces van de leerling gaat verder dan het zich herinneren van een kort antwoord.
- **Gebruik van significante inhoud**
De taak bevat enkel relevante aspecten van het te evalueren curriculum.
De taak integreert verschillende ideeën en concepten uit het curriculum.
De taak is rijk in die zin dat leerlingen worden aangemoedigd verder te denken dan het oplossen van de taak vereist.
- **Toegankelijkheid voor alle leerlingen**
Alle leerlingen kunnen de taak binnen een redelijke termijn aanvangen.
De taak maakt meerdere antwoordbenaderingen mogelijk.
De presentatie van de taak biedt verscheidene vormen van informatie (tekst, grafieken, tabellen, schema's, illustraties).
Er wordt een sobere taal gebruikt waarbij technische en onbekende termen worden vermeden.
- **Uitdagingheid**
De taak motiveert de leerlingen. Daartoe is ze uitdagend, stimulerend, origineel, nieuw en interessant.
De taak moedigt leerlingen aan zelfstandig beslissingen te nemen.
- **Gerichtheid**
De belangrijkste doelen van de taak zijn duidelijk en goed gedefinieerd.
De taak is zo opgesteld dat de leerlingen niet worden afgeleid van het centrale thema van de taak.
De taak begeleidt leerlingen, binnen bepaalde grenzen, bij het bereiken van de doelstellingen.

Deze lijst van karakteristieken mag dan wel overweldigend overkomen, toch is de stap naar dergelijke taken vaak niet zo moeilijk te zetten. Herinner je dat we reeds stelden dat de concepten van vaardigheidsevaluatie niet zo nieuw zijn. Het benadrukken van vaardigheidsevaluatie kan dus worden opgevat als een verderzetting of een herziening van reeds lang bestaande prak-

tijken (Mabry, 1999). Bovendien is het zo dat eenmaal leerkrachten de criteria voor prestatietaken onder de knie hebben, ze bedreven worden in het kiezen van geschikte evaluatietaken en in het ontwerpen en toepassen van de bijbehorende evaluatiecriteria.

4.2 Types vaardigheidstaken

Er zijn verschillende types vaardigheidstaken, gaande van eenvoudige ondervragingen en open vragen tot en met meer uitvoerigere types die van leerlingen verwachten dat ze hun prestaties presenteren of demonstreren (Mabry, 1999; Callison, 1998; Dochy, 1999b).

- ✓ Korte ondervragingen – open vragen. Vele leerkrachten gebruiken korte ondervragingen om te evalueren hoe goed hun leerlingen de noodzakelijke concepten en vaardigheden beheersen. Een korte ondervraging begint meestal met een zogenaamde stimulus (een wiskunde-probleem, een politieke cartoon, een kaart). Op basis van deze stimulus verwacht de leerkracht van de leerling dat hij gegevens interpreteert, beschrijft, berekeningen maakt, verklaart, voorspelt en antwoorden formuleert. De antwoorden vereisen vaak complexe en metacognitieve vaardigheden. De ondervraging kan zowel mondeling als schriftelijk gebeuren. Het antwoord van de leerling kan eveneens schriftelijk of mondeling zijn en kan bestaan uit een rekenkundige oplossing, een argumentatie, een tekening, diagram, grafiek, etc. Dit type prestatietak is niet nieuw, maar niettemin onvoldoende ingeburgerd.
- ✓ Projecten bestaan uit een gespecialiseerd, vaak interdisciplinair onderzoek georganiseerd en uitgevoerd door een leerling of een groep leerlingen. Van een project wordt verwacht dat het resulteert in gepersonaliseerde en nieuwe kennis, in ontwikkeling van competenties, in nagestreefde vaardigheden en in quasi beroepsmatige motivatie en handelwijzen. Ook van het samenwerken van leerlingen in projecten worden positieve resultaten verwacht.
- ✓ Demonstraties (of tentoonstellingen) zijn een eerder formele, meer of minder publiek toegankelijke demonstratie van de competenties en vaardigheden van de leerling die vooral gebruikt wordt als summatieve evaluatie. Demonstraties kunnen evenwel ook formatief en informeel worden ingepast in het curriculum en in de dagelijkse praktijk. Een demonstratie kan bijvoorbeeld worden ondersteund door tastbare producten van de leerling, resultaten van experimenten of oplossingen voor praktische problemen.
- ✓ Gespreksevaluatie is een evaluatie van wat de leerling zelf vertelt over wat hij/zij weet. Door te praten met een evaluator geeft de leerling aan wat hij heeft geleerd. De leerling staat dit met evidentie van kritische denkprocessen en probleemoplossend gedrag. Deze evidentie kan bestaan uit verhalen, anekdotes, argumenten, samenvattingen, interpretaties en zelfevaluatie. De leerling geeft als het ware een uiteenzetting over het eigen leerproces. De evaluator luistert naar de uiteenzetting van de leerling en tracht ook zelf gegevens te achterhalen die een beeld geven van de mate waarin de leerling het geleerde in nieuwe probleemsituaties kan toepassen.

- ✓ Portfolio's: een verzameling van informatie door en over de leerling die een breed beeld geeft van diens prestaties (aan dit type prestatietak besteden we aandacht in hoofdstuk 4 van dit deel).

4.3 De beoordeling van vaardigheidstaken

Het beoordelen van antwoorden en producten die het resultaat zijn van vaardigheidstaken moet zeer zorgvuldig gebeuren. Het unieke karakter van deze producten is immers dat de beoordeling afhankelijk is van de interpretatie van de evaluator. Dit maakt de beoordeling van vaardigheidstaken veel moeilijker dan de beoordeling van traditionele kennistoetsen (Janssens et al., 2000). Net zoals bij de andere alternatieve evaluatievormen die we zullen voorstellen, is het bij vaardigheidsevaluatie dan ook uiterst belangrijk de richtlijnen omtrent het ontwikkelen van criteria, rubrieken en evaluatielijsten na te leven. Prestatietaken kunnen de doelen die ze beogen te realiseren bovendien enkel waarmaken indien de beoordeling criteriumgeriefd is en indien voldoende tijd wordt uitgetrokken voor feedback. Bij het presenteren van een aantal voorbeelden van vaardigheidsevaluatie zal de beoordeling van vaardigheidstaken geïllustreerd worden.

5 Voorbeelden van vaardigheidstaken

Voor elk van de alternatieve evaluatievormen die we in dit deel presenteren, zullen we voorbeeldmateriaal ontwikkelen. Hiermee willen we in grote lijnen twee zaken bereiken: (1) het concrete gebruik van de betreffende alternatieve evaluatievorm illustreren en (2) aantonen dat het gebruik van alternatieve evaluatievormen geen ver-van-mijn-bed-show hoeft te zijn.

Een zoektocht in de evaluatieliteratuur levert al gauw een aantal mooie voorbeelden op van de verschillende evaluatievormen. Deze hier zonder meer overnemen zou -ons inziens- echter niet de meest gepaste werkwijze zijn. We bedoelen hiermee dat het niet volstaat om voorbeelden te kopiëren; we moeten aantonen dat het goed mogelijk is voor de eigen specifieke onderwijs- en evaluatiecontext concreet evaluatiemateriaal uit te werken. Vandaar dat we bij het ontwikkelen van voorbeelden in grote lijnen teruggrijpen naar de casussen die bij de aanvang van deel 2 beschreven hebben. Dit neemt niet weg dat we eerst een aantal niet uitgewerkte vaardigheidstaken presenteren. Op die manier wordt de variatie in mogelijke vaardigheidstaken duidelijk.

5.1 Algemene voorbeelden van vaardigheidstaken

Sommige vormen van vaardigheidstaken kunnen binnen verschillende vakken worden toegepast. Andere zijn eerder vakspecifiek. Van beide vormen geven we een aantal voorbeelden.

Figuur 4.5: Voorbeelden van vaardigheidstaken

Niet vakspecifiek

- ✓ Ontwerp een spel.
- ✓ Neem deel aan een debat.
- ✓ Gebruik een rubriek om een (eigen) prestatie te beoordelen.
- ✓ Ontwerp een museumbezoek.
- ✓ Verdedig een standpunt vanuit één invalshoek en veroordeel het vanuit een andere invalshoek.
- ✓ Maak een samenvatting van...
- ✓ Bekritiseer je eigen werk (geef voorbeelden en details).
- ✓ Welke argumenten ondersteunen de volgende stelling?
- ✓ Welke informatie heb je nodig om het volgende probleem op te lossen?
- ✓ Welke fouten zitten er in de volgende redenering?
- ✓ Als je deze gedachtegang verder zet, wat is dan de volgende stap?

Talen

- ✓ Herschrijf een kortverhaal in een toneelstuk.
- ✓ Schrijf een brochure om het gebruik van teveel antibiotica te ontraden.
- ✓ Schrijf een eigen einde aan een verhaal.
- ✓ Schrijf een aanbevelingsbrief voor je klasgenoot (of een fictief iemand).
- ✓ Leg uit welke problemen het personage tegenkomt en hoe hij ze oplost.
- ✓ Hervertel het verhaal in je eigen woorden.
- ✓ Op welke manier zou de tekst levendiger gemaakt kunnen worden?
- ✓ Vat in twee zinnen de boodschap van de auteur samen.

Wiskunde

- ✓ Los volgend vraagstuk op twee verschillende manieren op.
- ✓ Bewijs een stelling die nog niet in de les gezien is.
- ✓ Meet iets.
- ✓ Stel een begroting op.
- ✓ Zet een grafiektype om in een ander type.
- ✓ Stel volgend vraagstuk voor aan de hand van een vergelijking.

Wetenschappen

- ✓ Voer een experiment uit.
- ✓ Stel de data uit je experiment grafisch voor.
- ✓ Wat leert dit experiment over (het geleerde concept)?
- ✓ Hoe kan je op een veilige manier een labo inrichten?
- ✓ Hoe zou jij de wet van Faraday aan de hand van een experiment aantonen?
- ✓ Bekritiseer de opzet van volgend experiment.

5.2 Voorbeelden van vaardigheidstaken in onze casussen

Bij elke casus stellen we een aantal onderwijsdoelen voorop en presenteren we een vaardigheidstaak en een beoordelingsinstrument.

5.2.1. Vaardigheidsevaluatie bij leerlingen in de eerste graad Nederlands

Eén van de eindtermen Nederlands voor de leerlingen van de eerste graad luidt: de leerlingen kunnen voor hun leeftijd bestemde formulieren invullen. Daarnaast wordt van leerlingen verwacht dat ze een stuk met informatie over henzelf voor al of niet nader bekenden produceren. Deze schrijfvaardigheden kunnen geëvalueerd worden door leerlingen de volgende evaluatietask voor te leggen.

Figuur 4.6: Evaluatietask schrijfvaardigheden

Een man die onderzoek doet naar de veiligheid van de buurt waar je woont, vraagt je om de volgende vragenlijst in te vullen. Met die informatie zal de gemeente beslissen waar ze nieuwe veiligheidsmaatregelen moet nemen.

Vragenlijst (kleur de bolletjes en vul indien nodig verder aan)

Naam:

Straat:

Gemeente:

1. Hoe kom je naar school? 0 auto 0 fiets 0 bus 0 te voet 0

2. Vind je het veilig in jouw buurt? 0 ja 0 nee

Zo ja, waarom?

Zo nee, hoe zou jij de buurt veiliger maken?

3. Speel je vaak op straat? 0 ja 0 nee

Zo nee, waarom niet?

4. Woon je graag in deze buurt? Waarom (niet)?

5. Als je zelf mocht kiezen, in welke buurt zou je dan willen wonen?

Om de leerlingenprestaties te beoordelen, doen we een beroep op een analytische, niet-taakgebonden rubriek met vier criteria en vier prestatieniveaus.

Figuur 4.7: Rubriek voor het invullen van formulieren

	Excellent	Goed	Voldoende	Moet verbeteren
Verstaan van verschillende vraagtypes	De leerling begrijpt duidelijk de structuur van de vragenlijst en de samenhang tussen de vragen.	De leerling begrijpt de structuur van de vragenlijst en de samenhang tussen de vragen.	De leerling heeft een enkele keer de structuur en de samenhang tussen de vragen niet door.	De structuur en de samenhang tussen de verschillende vragen is onvoldoende tot de leerling doorgedrongen.
Analyseren van de vraag	De leerling verstaat zeer duidelijk naar welke informatie de vragen peilen.	De leerling verstaat naar welke informatie de vraag peilt.	De leerling verstaat bij de meeste vragen naar welke informatie gepeild wordt.	De leerling verstaat bij de meeste vragen niet naar welke informatie gepeild wordt.
Informatie	Alle informatie is relevant, duidelijk en grondig.	De meeste informatie is relevant, duidelijk en grondig.	De meeste informatie is duidelijk en relevant maar niet altijd voldoende grondig.	De informatie is vaak niet relevant of onduidelijk.

Netheid	De vragenlijst is zeer ordelijk en net ingevuld zonder doorhalingen.	De vragenlijst is ordelijk en net ingevuld met één of twee doorhalingen.	De vragenlijst is leesbaar en sommige delen zijn verzorgd. Andere delen lijken te snel ingevuld.	De vragenlijst is niet ordelijk en net ingevuld. De leerling lijkt niet begaan met de netheid ervan.
---------	--	--	--	--

Ook spreekvaardigheden komen aan bod in de eindtermen. Men verwacht bijvoorbeeld van leerlingen dat ze instructies kunnen geven aan bekende leeftijdsgenoten, dat ze in een telefoongesprek onbekende volwassenen informatie kunnen geven en dat ze een eigen standpunt onder woorden kunnen brengen. Eerst presenteren we voor elke doelstelling een vaardigheidstaak en daarna stellen we een holistische rubriek samen die kan gebruikt worden bij het beoordelen van de leerlingprestaties op elk van deze taken.

Figuur 4.8: Evaluatietaken voor spreekvaardigheden

Instructies geven. Denk even aan iets dat je zelf kan koken of klaarmaken (een spiegelei, een pizza of een volledig middagmaal). Leg aan je medeleerling uit hoe hij/zij stap voor stap te werk moet gaan om zelf dat gerecht klaar te maken.

Informatie geven. Stel dat je alleen thuis bent. Op een gegeven moment ruik je een brandgeur. Je belt de brandweer en je krijgt mij aan de lijn. Vertel mij alles wat ik moet weten om jou te kunnen helpen.

Een eigen standpunt onder woorden brengen. Vind jij dat wij in onze klas en school voldoende inspanningen doen om ons afval te sorteren? Vertel jouw mening aan de klasgenoten.

Figuur 4.9: Holistische rubriek voor spreekvaardigheid

	Criteria	Bekwaam spreekniveau	Aanvaardbaar spreekniveau	onvoldoende spreekniveau
Overdracht	volume snelheid articulatie	De drie criteria zijn goed en minstens twee zijn zeer goed.	De drie criteria zijn goed.	Minstens 1 criterium is onvoldoende.
Organisatie	volgorde verbanden	De twee criteria zijn zeer goed.	De twee criteria zijn goed.	Minstens 1 criterium is onvoldoende.
Inhoud	hoeveelheid relevantie aangepastheid	De drie criteria zijn goed en minstens twee zijn zeer goed.	De drie criteria zijn goed.	Minstens 1 criterium is onvoldoende.
Taal	zinsbouw woordkeuze	De twee criteria zijn zeer goed.	De twee criteria zijn goed.	Minstens 1 criterium is onvoldoende.

5.2.2. Vaardigheidsevaluatie bij leerlingen in de eerste graad wiskunde

Voor leerlingen in de eerste graad van het secundair onderwijs worden in de eindtermen en leerplannen onder andere volgende wiskundedoelen vooropgesteld: (1) de leerlingen voeren de hoofdbewerkingen correct uit in de verzamelingen van de natuurlijke, de gehele en de rationale getallen en (2) de leerlingen kunnen een goede schets of een aangepast schema van een wiskundig probleem maken.

Figuur 4.10: Een rekentaak

Francis en Eva hebben een huisje in de vorm van een vierkant. Het huis meet 6m op 6m. Op een beurs voor bouwmaterialen wonnen ze de hoofdprijs. Deze bestaat uit 80 m² terrastegels in een kleur naar keuze. Morgen gaan Francis en Eva beslissen in welke vorm ze het terras gaan aanleggen. Ze willen met de tegels een pad rond het huis aanleggen. Dat pad moet tegen het huis beginnen. Ze vragen zich af of ze, als ze een pad van 5 m breed aanleggen, zullen toekomen met de tegels die ze gewonnen hebben. Kan jij Francis en Eva even helpen met het beantwoorden van deze vraag? Geef in de oplossing alle wiskundige bewerkingen weer die jouw antwoord ondersteunen. Maak ook een schema of een schets van het probleem.

Stel dat we een algemene beoordeling willen uitspreken over de prestaties van de leerlingen dan kunnen we een beroep doen om een holistische rubriek. We stellen de rubriek zo samen dat hij ook nog bij het beoordelen van andere antwoorden op dergelijke vragen dienst kan doen. Hierbij onderscheiden we vier prestatieniveaus.

Figuur 4.11: Een holistische rubriek voor het oplossen van rekenproblemen

4-Een voorbeeldoplossing. De leerling geeft een volledig, duidelijk en coherent antwoord. De berekeningen zijn allemaal correct. De leerling beheerst de verschillende bewerkingen. Het schema is een duidelijke weergave van het probleem. De leerling identificeert al de relevante elementen van het probleem.

3-Een goede oplossing. De leerling geeft een duidelijk en coherent antwoord. Een aantal details ontbreekt. De berekeningen zijn correct. Het schema is een duidelijke weergave van het probleem maar bevat niet alle gegevens. De leerling identificeert de belangrijkste elementen van het probleem.

2-Een voldoende oplossing. Het antwoord is voldoende maar de uitleg schiet soms tekort. Het schema is enigszins onduidelijk maar er blijkt wel uit dat de leerling de essentie van het probleem begrijpt. Een enkele rekenfout maakt dat het antwoord foutief is. Fouten zijn toevallig en worden niet systematisch gemaakt. De leerling besteedt in het antwoord aandacht aan een aantal niet-relevante elementen.

1-Een onvoldoende prestatie. Het antwoord is onduidelijk of is afwezig. De leerling slaagt er niet in het probleem op een duidelijke manier in het schema weer te geven. De essentie van het probleem is niet begrepen. Er zijn veel rekenfouten.

Alvorens leerlingen te kunnen beoordelen op een dergelijke holistische schaal moeten evaluatiegegevens op een rij worden gezet. Dat kan bijvoorbeeld aan de hand van volgende evaluatielijst.

Figuur 4.12: Een evaluatielijst voor het oplossen van rekenproblemen

	4	3	2	1
De leerling slaagt er vlot in open rekenproblemen aan te pakken.	O	O	O	O
De leerling geeft een duidelijk geformuleerd antwoord.	O	O	O	O
De leerling geeft al de noodzakelijk bewerkingen weer.	O	O	O	O
De leerling respecteert de volgorde van de bewerkingen.	O	O	O	O
De leerling maakt correcte berekeningen.	O	O	O	O
Het schema stelt het probleem duidelijk voor.	O	O	O	O

Het schema bevat al de noodzakelijk gegevens.	0	0	0	0
De leerling controleert het antwoord.	0	0	0	0
...	0	0	0	0
1= voorbeeldig; 2= goed; 3=voldoende; 4= onvoldoende				

5.2.3. Vaardigheidsevaluatie bij leerlingen Elektriciteit/Elektronica

Indien we willen evalueren of leerlingen op een gestructureerde manier fouten in een signalisatieschakeling (zoals een bel) kunnen opzoeken, volstaat het niet om de leerlingen de verschillende stappen die ze zullen ondernemen te laten opsommen. De leerlingen moeten de verschillende stappen niet alleen kennen, ze moeten ze ook kunnen toepassen in nieuwe situaties. Dat kunnen ze bijvoorbeeld aantonen aan de hand van de volgende evaluatietask tijdens een mondeling evaluatiemoment.

Figuur 4.13: Storingen in signalisatie-installaties opsporen

Achtergrond. Leerlingen krijgen de opdracht de storing op te sporen in een opstelling van een belinstallatie. De oproepbel werkt niet meer. De oorzaak hiervan is niet direct zichtbaar vermits het onderbrekingscontact in de bel is geoxideerd. Leerlingen kunnen van alle relevante meetinstrumenten gebruik maken en kunnen aan de leerkracht vragen stellen om bijkomende inlichtingen in te winnen over de storing.

De leerlingen moeten mondeling hun gedachtegang verwoorden en verantwoorden.

De evaluatietask. Hoe kan je op de meest efficiënte manier de foutoorzaak van het niet werken van deze bel opsporen.

Om de activiteiten van de leerling te beoordelen, maakt de leerkracht gebruik van een checklist (zie figuur 4.15). Er zijn vier belangrijke stappen in het opsporen van fouten. De beoordeling op deze vier criteria is gebaseerd op de beoordeling van een aantal deelaspecten. Aan de hand van de beoordeling op de verschillende items in de checklist wordt de leerling vervolgens een algemene score toegekend. Hiertoe kan de holistische rubriek in figuur 4.14 gebruikt worden.

Figuur 4.14: Holistische rubriek voor het opsporen van fouten in een signalisatie-installatie

Goed. De leerling vindt aan de hand van gepaste vragen relevante informatie in over de storing. Wanneer hij over voldoende informatie beschikt, stelt de leerling het probleem zelf vast. Om de fout te lokaliseren, verwerft de leerling gepaste informatie over het probleem. Hierbij gaat hij op een georganiseerde manier tewerk. De leerling lokaliseert de fout door met aangepaste handelingen en metingen het aantal mogelijke fouten te verminderen en een bevestiging te krijgen van de veronderstelde fout.

Voldoende. De leerling vindt informatie in over de storing en stelt zelf het probleem vast. Om de fout te lokaliseren, verwerft de leerling gepaste informatie over het probleem. Dit gebeurt evenwel niet op de meest efficiënte manier. De leerling lokaliseert de fout maar het verwerven van inzicht in het probleem en de eigenlijke lokalisatie van de fout verlopen onvoldoende georganiseerd om van een goede prestatie te kunnen spreken.

Onvoldoende. De leerling vindt onvoldoende of irrelevante informatie in over de storing. De verschillende stappen in het opsporen van storingen in signalisatie-installaties verlopen niet in een logische volgorde. De oorzaak van het probleem wordt lukraak gelokaliseerd, de leerling heeft geen systematisch aanpak. De fout wordt ofwel niet, ofwel door toeval of na zeer lang zoeken vastgesteld.

Figuur 4.15: checklist opsporen van fouten in een signalisatie-installatie

	G	V	O
Inlichtingen inwinnen over de storing	0	0	0
Gericht vragen stellen over de geschiedenis van het probleem	0	0	0
Duidelijkheid van de vragen voor leken	0	0	0
Logica in de gestelde vragen	0	0	0
Zelf de fout vaststellen (door meting of gebruik)	0	0	0
Inzicht verwerven in het probleem	0	0	0
Het schema van de installatie bestuderen	0	0	0
Nagaan hoe de bel normaliter bediend moet worden	0	0	0
Mogelijke foutoorzaken identificeren	0	0	0
Voeding van de installatie controleren	0	0	0
Werking van de schakelaar controleren	0	0	0
De fout lokaliseren	0	0	0
Op een logische manier mogelijke oorzaken uitsluiten	0	0	0
Op een gepaste manier controleren of de gestelde diagnose klopt	0	0	0
<i>Beoordelingsschaal: Goed, Voldoende, Onvoldoende</i>			

5.2.4. Vaardigheidsevaluatie bij leerlingen Verzorging

Voor de derde graad van het secundair onderwijs hebben we als casus de leerlingen BSO binnen de richting Verzorging vooropgesteld. Van deze leerlingen wordt verwacht dat ze onder andere volgende doelen bereiken: leerlingen kunnen de begrippen 'afhankelijkheid', 'zich afhankelijk voelen' en 'afhankelijk gedrag' in hun eigen woorden beschrijven; leerlingen kunnen in het gedrag van ouderen vormen van afhankelijkheid herkennen en er gepast op reageren. Om deze doelen te evalueren, kunnen we onder meer een beroep doen op de volgende vaardigheidstask.

Evaluatietask: Lees eerst even de volgende tekst over twee vrouwen in een rusthuis (Verhoef & Visser, 1994). Beantwoord daarna de onderstaande vraag.

Mevrouw Peters is halfzijdig verlamd en zit in een rolstoel. Dat maakt dat zij de hele dag en in verschillende opzichten afhankelijk is van een verzorgende. Dat vindt mevrouw Peters vreselijk. Bijvoorbeeld wanneer ze op anderen moet wachten om uit bed te komen of om naar het toilet te gaan. Het liefst zou zij alles nog zelfstandig doen. Mevrouw Lenaerts is evenzeer afhankelijk. Zij is eigenlijk nog goed te been, maar vraagt toch bij alles wat ze doet de assistentie van het verzorgend personeel.

De opdracht

Stel dat jij de stagebegeleidster bent van een leerling uit de studierichting Verzorging. De leerling vraagt je hoe ze op beide zorgvragers moet reageren. Ze vraagt zich af of ze beide bejaarden dezelfde steun en hulp moet geven?

Wat zou jij haar adviseren? Geef hieronder zowel aan wat de leerling zou moeten doen (aandachtspunten) als waarom ze zo moet handelen.

Als we geen rekening houden met de schrijfvaardigheden van leerlingen kunnen we in deze taak de volgende drie leeruitkomsten (of criteria) onderscheiden: het gebruik van begrippen; het herkennen van voorbeelden van 'afhankelijkheid', 'zich afhankelijk voelen' en 'afhankelijk gedrag'; en het omgaan met afhankelijk gedrag. Om de antwoorden van leerlingen te beoordelen, kunnen we een beroep doen op de volgende analytische rubriek.

Figuur 4.16: Rubriek omtrent het begrip afhankelijkheid

	Goed	Voldoende met leentes	Onvoldoende
Gebruik van begrippen	De leerling gebruikt het begrip-paar 'afhankelijk zijn' en 'zich afhankelijk gedragen' correct. Het onderscheid wordt duidelijk uitgelegd.	De leerling gebruikt het onder-scheid tussen afhankelijk zijn en zich afhankelijk gedragen maar verwijst niet naar het juiste begrippen-paar.	De leerling heeft onvol-doende inzicht in het onderscheid tussen afhankelijk zijn en zich afhankelijk gedragen.
Herkennen van voor-beelden	De leerling haalt uit de casus en uit de eigen (stage) ervaringen goede voorbeelden aan van de verschillende begrippen.	De leerling haalt goede voor-beelden aan maar brengt deze soms onder bij de verkeerde begrippen.	De voorbeelden die de leerling aanhaalt, sluiten niet aan bij de verschil-lende begrippen.
Omgaan met afhanke-lijkheid	De leerling geeft duidelijk aan dat bij afhankelijk zijn en zich afhankelijk gedragen op ver-schillende manieren moet gereaa-geerd worden. De leerling zet daarnaast de belangrijkste aan-dachtspunten in het omgaan met afhankelijk gedrag op een rij.	De leerling geeft aan dat bij afhankelijk zijn en zich afhanke-lijk gedragen op verschillende manieren gereageerd moet wor-den. Er worden onvoldoende of minder relevante aandachtspun-ten gegeven om met afhankelijk gedrag om te gaan.	De leerling geeft niet aan dat op verschillende manieren moet gereaa-geerd worden. Er wor-den geen of niet rele-vante aandachtspunten gegeven.

Hoofdstuk 2: Authentieke evaluatie

Authentieke evaluatie wordt vaak in één adem genoemd met vaardigheidsevaluatie. Dat komt omdat beide evaluatiebenaderingen gemeenschappelijk hebben dat leerlingen een antwoord of een product moeten construeren. Authentieke evaluatie is dan ook een vorm van vaardigheidsevaluatie. Kenmerkend is evenwel dat een bijkomend kenmerk van alternatieve evaluatievormen wordt nagestreefd, namelijk de authenticiteit (levensechtheid of realiteit) van evaluatietaken. De benadrukte kenmerken van alternatieve evaluatie zijn dus zowel levensechtheid als constructiegerichtheid. In dit hoofdstuk besteden we voornamelijk aandacht aan de levensechtheid. De constructiegerichtheid werd in het vorige hoofdstuk immers uit de doeken gedaan.



Figuur 4.17: Authentieke evaluatie als alternatieve evaluatievorm

1 Authentieke evaluatie geïllustreerd

Authentieke evaluatie kan in bepaalde domeinen gemakkelijker worden aangewend dan in andere. De meest bekende voorbeelden situeren zich in het hoger onderwijs. Domeinen die zich daar zeer goed lenen tot het gebruik van authentieke evaluatietaken zijn bijvoorbeeld opleidingen geneeskunde en onderwijzer. De medische vaardigheden van studenten kunnen bijvoorbeeld levensecht worden geëvalueerd tijdens een rollenspel met gestandaardiseerde patiënten. Dat zijn personen die getraind zijn om verschillende ziektes te simuleren. Studenten in de lerarenopleiding kunnen geëvalueerd worden in een echt leslokaal met echte leerlingen.

Ook in het secundair onderwijs kunnen we mooie illustraties vinden van situaties waarin authentieke evaluatietaken gebruikt worden. Binnen het vak Nederlands bijvoorbeeld: een belang-

rijk doel is daar de verschillende aspecten van taalvaardigheid te evalueren in contexten die nauw aansluiten bij de eigenlijke situaties waarin de schrijf- en leesvaardigheden worden gebruikt. Zo zullen authentieke evaluatietaken van leerlingen verwachten dat ze krantenartikelen lezen, dat ze over betekenisvolle en levensechte inhoud schrijven en dat ze taken uitvoeren zoals het bespreken van boeken en gedichten, en het samenstellen van brochures en tijdschriften. Ook andere vakken lenen zich tot het gebruik van authentieke leer- en evaluatietaken. Hierbij denken we aan het uitvoeren van wetenschappelijke experimenten, van onderzoek en dergelijke. Toch hoeven levensechte taken niet zo uitgebreid en verregaand te zijn. Ook het oplossen van rekenvraagstukken die toepassingen kennen in het dagelijks leven verdient immers de term alternatieve evaluatievorm.

2 Authentieke evaluatie gedefinieerd

We stelden reeds dat authentieke evaluatie een vorm is van vaardigheidsevaluatie. Authentieke evaluatievormen zijn met andere woorden een deelverzameling van constructiegericht evaluatie¹. Net zoals vaardigheidstaken verwachten authentieke evaluatietaken dat leerlingen aantonen dat ze verworven kennis effectief kunnen gebruiken, dat ze antwoorden construeren. Het onderscheid tussen beide evaluatievormen zit in de realistische context waarin de evaluatietaken worden aangeboden en uitgevoerd.

Wij hanteren de volgende definitie van authentieke evaluatie(taken) (gebaseerd op Mabry, 1999; Wiggings, 1990; Callison, 1998):

Authentieke evaluatie beoordeelt de mogelijkheden van leerlingen aan de hand van realistische taken en situaties. Authentieke evaluatietaken verwachten van leerlingen dat ze vaardigheden en competenties aantonen op een manier waarop professionals deze vaardigheden en competenties in de dagelijkse realiteit aanwenden. We spreken dus van authentieke evaluatie wanneer de evaluatietaken gelijken op de taken die professionals ondernemen of -ruimer opgevat- wanneer de taken authentiek zijn in die zin dat ze de zogenaamde realistische wereld buiten de school representeren.

Een evaluatie is authentiek indien ze op zichzelf betekenisvol is, indien de leerling de taak en de context ervan als levensecht ervaart en indien de prestaties die geëvalueerd worden ook betekenisvol zijn buiten het klaslokaal (Kerka, 1995).

¹ Sommige auteurs stellen de levensechtheid tevens als een kenmerk van vaardigheidstaken voorop (WCER, 1993). In dat geval wordt er dus niet echt een onderscheid tussen beide evaluatievormen gemaakt en wordt aan de lijst van kenmerken van vaardigheidstaken die we in het vorige hoofdstuk presenteerden het element authenticiteit toegevoegd. Anderen baseren het onderscheid enkel op het feit dat authentieke evaluatie meestal meer tijd vergt dan vaardigheidsevaluatie.

3 Kenmerken en doelen van authentieke evaluatie

Authentieke taken moeten tevens constructiegericht zijn. Ze voegen een kenmerk toe aan de vaardigheidstaken. We volstaan hier dus met het beschrijven van de kenmerken en doelen van dit ene kenmerk dat maakt dat authentieke evaluatie nog meer dan constructiegerichte taken de term alternatieve evaluatievormen verdienen, namelijk de levensechtheid van de evaluatietaken.

3.1 Kenmerken van authentieke evaluatie

De authenticiteit van een evaluatievorm is afhankelijk van een drietal factoren: de evaluatietaak zelf, de context waarin de taak moet worden uitgevoerd en de gehanteerde evaluatiecriteria. Elk van deze elementen moet ertoe bijdragen dat de evaluatiesetting zo getrouw mogelijk de realistische context benadert zoals deze door professionals ervaren wordt. Dit betekent dat zowel het materiaal dat gebruikt wordt om de taak te omkaderen als de taak zelf zo natuurlijk mogelijk moeten zijn.

Net zoals de andere kenmerken van alternatieve evaluatievormen moet levensechtheid opgevat worden als een continuüm. Ook levensechtheid is geen kwestie van alles of niets. Het komt erop neer een zo realistisch mogelijke evaluatie te ontwerpen. De grens tussen authentieke evaluatie en niet authentieke evaluatie is moeilijk vast te leggen. Authenticiteit is bovendien een relatief begrip. Het is niet realistisch van een evaluatie te verwachten dat ze volledig authentiek zou zijn (Wiggings, 1990). Te denken valt hierbij bijvoorbeeld aan het afleggen van een rijexamen. Ook al is een dergelijke praktische rijtest veel authentiekere dan een 'paper and pencil'-test, het blijft een feit dat een examinerator naast de student zit. In die zin blijft de evaluatietaak toch afwijken van de situatie van de normale chauffeur wanneer die zich in het verkeer begeeft. Daartegenover staat echter dat het niet volledig kunnen realiseren van authentieke contexten ook duidelijke voordelen kan hebben. Leerlingen kunnen immers in die omstandigheden immers risico's nemen, experimenteren en leren uit hun fouten.

3.2 Doelen van authentieke evaluatie

Het doel van authentieke evaluatie kent zijn oorsprong in kritieken op traditionele evaluatievormen. Al te vaak moest men immers constateren dat deze evaluatievormen niet de vaardigheden en bekwaamheden evalueren die leerlingen in het dagelijks leven nodig hebben om succesvol te zijn. De bekwaamheden en vaardigheden die er echt toe doen, zijn te complex om via de geïsoleerde componenten van traditionele toetsen te evalueren. Alternatieven dringen zich dus op. Als het doel van de evaluatie is te achterhalen of leerlingen kunnen functioneren buiten de school, is het aangewezen een beroep te doen op authentieke evaluatie. Evaluaties die geen betekenis hebben in niet-schoolse contexten, zoals het toetsen van woordenschat en historische data, leveren met dat doel voor ogen geen informatieve evaluatiegegevens op. In het echte leven gebruiken we woorden bij ons spreken en schrijven en worden we zelden verondersteld ons

historische data te herinneren. Het komt er dus op aan taken te ontwikkelen die indicaties kunnen geven voor de vraag of leerlingen effectief kunnen presteren in realistische situaties. Hierin schuilt dan ook het belangrijkste doel van het gebruik van authentieke evaluatie: nagaan of leerlingen blijf geven van een succesvolle beheersing van kennis, vaardigheden en attitudes in realistische situaties waartoe ze worden opgeleid en opgevoed. De nadruk ligt hierbij vaak op het evalueren van de creativiteit van leerlingen bij het oplossen van problemen, het plannen van activiteiten, de mogelijkheden om kennis te integreren en het samenwerken met anderen (MMS, 2001).

Werken aan authentieke taken is een zinvolle, uitdagende activiteit en het verzekert dat men het onderwijsleerproces richt op waardevolle vaardigheden. Bovendien maken het realistische karakter van de taken en het gegeven dat leerlingen weten wat er van hen verwacht wordt dat het prestatieniveau van alle leerlingen tot hogere kwaliteitsniveaus gebracht kan worden (Wiggins, 1990).

Naast de voordelen, nastrevenswaardige doelen en mogelijkheden die het gebruik van authentieke evaluatietaken kan hebben, moeten eveneens een aantal mogelijke nadelen aangehaald worden. Zo is het bijvoorbeeld -op basis van de geobserveerde leerlingprestaties- moeilijk te achterhalen wat moet worden toegeschreven aan de voorkennis dan wel aan vaardigheden die de leerlingen doorheen het onderwijsleerproces ontwikkeld hebben. Verder is het wellicht zo dat het uitdagende karakter van de taken niet bij alle leerlingen motiverend werkt. Wanneer leerlingen als onderdeel van een taak in groep participeren, kan het bovendien moeilijk zijn de individuele prestaties van alle leerlingen te beoordelen.

4 Authentieke evaluatietaken

In wezen is het omschrijven van authentieke evaluatietaken eenvoudig: een authentieke evaluatietaken kenmerkt zich door het feit dat men zoveel mogelijk realisme nastreeft. Om dergelijke realistische taken te ontwerpen, is het aangewezen de volgende karakteristieken van authentieke taken in acht te nemen (Mabry, 1999; Wiggins, 1990; MMS, 2001; Kerka, 1995; Callison, 1998).

Figuur 4.18: Een aantal karakteristieken van authentieke taken

- ✓ Authentieke taken en evaluatiesettings benaderen zo getrouw mogelijk de realistische context.
- ✓ Taken worden niet alleen uitgevoerd omwille van het gebruiken van geleerde symbolen, regels en procedures. Indien mogelijk heeft de taak dus ook een levensecht doel. Dat maakt de taak betekenisvol en uitdagend.
- ✓ De taken maken, waar mogelijk, gebruik van informatie- en communicatietechnologieën.
- ✓ Leerlingen hebben voldoende tijd om de taak of opdracht te plannen.
- ✓ Leerlingen mogen alternatieve oplossingen aanreiken. Er is meer dan één correct antwoord.
- ✓ Leerlingen krijgen de mogelijkheid zichzelf te evalueren.
- ✓ Leerlingen hebben de mogelijkheid hun prestaties met anderen te evalueren en om advies te vragen. Werken in groep en individuele activiteiten worden afgewisseld.

Authentieke taken kunnen allerlei vormen aannemen zodat het moeilijk is er een aantal groepen in te onderscheiden. In principe zouden we kunnen verwijzen naar de groepen prestatietaken die bij het uitwerken van vaardigheidsevaluatie werden besproken. Indien deze vaardigheidstaken tevens een levensecht karakter krijgen, is er immers sprake van authentieke taken. Ook authentieke evaluatietaken kunnen dus variëren van korte open vragen tot heuse projecten.

Het onderscheid tussen een vaardigheidstaak en een authentieke taak is ondertussen duidelijk. Maar hoe kan je authentieke taken ontwikkelen, waar kan je inspiratie vinden en hoe zien die taken er uit? De volgende paragrafen gaan op deze vragen in. Om een specifiek beeld te geven van authentieke evaluatietaken zetten we eerst een aantal kenmerken van evaluatietaken op een rij. Op die manier kunnen we het onderscheid met vaardigheidstaken scherper stellen. Daarna bespreken we een aantal inspiratiebronnen en typische vormen van authentieke taken.

4.1 Hoe vaardigheidstaken authentiek maken?

Welke extra's maken van vaardigheidstaken authentieke taken? Net zoals de andere kenmerken van alternatieve evaluatievormen is authenticiteit een continuüm. Het komt er op aan vaardigheidstaken zo authentiek mogelijk te maken. De volgende lijst geeft een aantal ideeën en werkwijzen die kunnen bijdragen tot een authentieke evaluatiecontext.

Figuur 4.19: Komen tot authentieke evaluatietaken

- ✓ *Levensechtheid.* De leerlingen moeten de link tussen wat geëvalueerd wordt en de buitenschoolse wereld begrijpen. Leerlingen Personenzorg kunnen ter beoordeling van hun vaardigheden om actief te luisteren een gesprek met een bejaarde op tape vastleggen. Leerlingen Elektriciteit/Elektronica kunnen ter beoordeling van hun kennis van en attitudes tegenover veiligheidsvoorschriften bijvoorbeeld een brochure schrijven die de relevante veiligheidsvoorschriften met een aantal handige tips op een rij zet voor een specifieke werksituatie (uitslijpen elektriciteit, aansluiten elektriciteitskast, inbouwen elektrische toestellen).
- ✓ *Samenwerken.* Leerlingen worden verwacht actief te participeren aan het realiseren van projecten en te functioneren in coöperatieve groepen. De leerlingen helpen mekaar en worden niet tegen mekaar afgewogen bij de beoordeling. Door in groep te werken en taken te verdelen kunnen meer uitgebreide opdrachten worden uitgevoerd en kunnen leerlingen van mekaar leren. Leerlingen die binnen het vak Nederlands een schoolblad uitgeven, kunnen bijvoorbeeld afwisselend verschillende tekstsoorten maken (opiniërend, verhalend, informatief, humoristisch). Door in groepjes fouten in het dossier van een elektrische installatie op te sporen leren leerlingen Elektriciteit in kleine professionele teams samenwerken, plannen, taken verdelen en coördineren.
- ✓ *Communicatie en informatie.* De leerlingen hebben toegang tot verscheidene informatiebronnen en communicatietechnieken. Om de taak op te lossen, moeten leerlingen kunnen beschikken over de informatiebronnen waarover een professional in het betreffende werktein beschikbaar is. Leerlingen kunnen tevens op verschillende manieren met mekaar en met de leerkracht in contact komen.

- ✓ *Verantwoordelijkheid.* De leerlingen worden niet sterk gestuurd bij het uitvoeren van de evaluatietaken. Waar mogelijk kunnen leerlingen zelfstandig keuzes maken en hun oplossingsstrategie plannen. Ook de manier waarop leerlingen hun oplossing rapporteren en eventueel presenteren, valt onder hun eigen verantwoordelijkheid.
- ✓ *Flexibiliteit.* Het tijdschema biedt leerlingen voldoende flexibiliteit om hun taak of project succesvol te beëindigen.

4.2 Inspiratiebronnen voor authentieke taken

De korte argumentatie voor en de beschrijving van het gebruik van authentieke evaluatietaken maken nog niet volledig duidelijk hoe dergelijke evaluatietaken er kunnen uitzien. Vandaar dat we hier aandacht besteden aan een aantal voorbeelden. Aan de hand van deze voorbeelden willen we tevens illustreren waar leerkrachten ideeën kunnen halen voor het creëren van authentieke en constructiegerichte prestatietaken (ASCD, 1996). We onderscheiden de volgende bronnen.

Figuur 4.20: Bronnen voor het ontwikkelen van authentieke en constructiegerichte taken

- ✓ *Teksten in schoolboeken.* Stel dat een hoofdstuk in een aardrijkskundeboek is gewijd aan de veelheid van landbouwproducten, mineralen en technologieën die de verschillende gemeenschappen in België rijk zijn. Deze kennis kan op verschillende manieren geëvalueerd worden. Een mogelijk authentieke taak bestaat erin de leerlingen aan de hand van de tekorten van de andere gemeenschappen een promotieplan te laten opstellen dat vertegenwoordigers van de andere gemeenschappen informeert over welke producten en technologieën de eigen gemeenschap hen als surplus kan aanreiken. Of men kan de leerlingen de tekst in het schoolboek laten herschrijven en illustreren voor een jongere leeftijdsgroep.
- ✓ *Het curriculum.* (1) Stel dat het curriculum vooropstelt dat alle leerlingen een opiniërend essay moeten kunnen schrijven. Waarom leerlingen in dat geval geen standpunt laten innemen in een voor hen actuele discussie. Hun taak bestaat er dan bijvoorbeeld in aan het schoolbestuur of lokale politici een artikel te schrijven. (2) Stel dat het curriculum stelt dat leerlingen moeten weten hoe mensen zich tegen verschillende weertypes kunnen beschermen. Waarom leerlingen in dat geval geen brochure laten maken die uiteenzet welke weertypes gevaarlijk zijn en hoe men zich er tegen kan beschermen?
- ✓ *Actualiteit.* Door een oog op het nieuws te richten, wordt het als leerkracht mogelijk de relevantie voor wat de leerlingen leren voor het buitenschoolse leven aan te geven. Actualiteit kan ook inhouden dat men zorgt voor een effectief publiek voor de resultaten van de leertaken.
- ✓ *Authetieke tekstmaterialen.* Bepaalde werken (fictie en non-fictie) kunnen een interessante invalshoek vormen voor constructiegerichte en authentieke (evaluatie)taken. Te denken valt bijvoorbeeld aan het relateren van George Orwell's 'Animal Farm' aan hedendaagse politieke ontwikkelingen en aan de taak waarin de leerlingen fungeren als jury voor het selecteren van zowel interessante als wetenschappelijk correcte kinderboeken. In ruime zin kunnen authentieke tekstmaterialen evenwel nog op verschillende andere manieren inspirerend zijn (reisbrochures, kookboeken, catalogi, reclamefolders, menukaarten etc.).

5 Voorbeelden van authentieke taken

Ook de levensechtheid van een evaluatietaken moet worden opgevat als een continuüm. Dat betekent dat het vaak moeilijk is een duidelijk onderscheid te maken tussen vaardigheidstaken en authentieke taken. Op zich is het ook niet zo belangrijk om het onderscheid te kunnen maken. Wat wel van belang is, is dat de levensechtheid voldoende wordt nagestreefd. Dat kan bijvoorbeeld aan de hand van de volgende evaluatietaken en beoordelingsinstrumenten.

5.1 Algemene voorbeelden van authentieke taken

Authentieke taken voegen een kenmerk toe aan de vaardigheidstaken die we in het vorige hoofdstuk presenteerden. Mits het maken van een aantal aanpassingen aan de inhoud en de context kunnen de voorbeelden van vaardigheidstaken worden omgevormd tot levensechte taken. Bijvoorbeeld: in plaats van enkel een spel te ontwerpen gaan we het spel ook effectief spelen; in plaats van zo maar teksten te schrijven, worden deze gebundeld in een tijdschrift voor een authentiek publiek; in plaats van rekenopdrachten op te lossen, stellen de leerlingen een begroting op voor hun eigen activiteiten tijdens het schoolfeest. De levensechtheid situeert zich dus niet alleen in de aard van de taak, maar ook in de inhoud en de context. Twee zaken staan hierbij telkens centraal: (1) verzamel je met de taak relevante informatie over de vraag of leerlingen de vooropgestelde kennis en vaardigheden kunnen gebruiken in binnenschoolse en buitenschoolse (beroepsmatige) situaties en (2) heeft de taak een doel op zich, is het geen taak om de taak?

5.2 Voorbeelden van authentieke taken in onze casussen

De voorbeelden van vaardigheidstaken voor de casussen die we selecteerden, zijn in meer of mindere mate levensecht. In een aantal gevallen kan de levensechtheid van de taken door inhoud en context te variëren, nog vergroot worden. Zo zou je de leerlingen realistische formuleren kunnen laten invullen en deze ook effectief insturen (om lid te worden van een boekenclub, om zich kandidaat te stellen voor een vakantiejob of voor een plaats in het schoolsportteam). Je kan leerlingen een standpunt laten verwoorden over iets dat later nog aan bod zal komen. Of je kan hun standpunten doorspelen naar andere betrokkenen. Leerlingen kunnen berekenen hoeveel liter verf er nodig is om in de zomer hun klaslokaal te laten schilderen. In sommige gevallen kan je leerlingen betrekken in levensechte situaties. Dat kan tijdens de stages, maar ook indirect. Dat betekent bijvoorbeeld dat leerlingen levensechte elektriciteitsplannen bestuderen en eventueel de uitvoering ervan ter plaatse gaan controleren.

In de volgende paragrafen geven we voor de verschillende casussen telkens een nieuw voorbeeld van een authentieke evaluatietaken met bijbehorende beoordelingsinstrument.

5.2.1. Authentieke evaluatie bij leerlingen in de eerste graad Nederlands

Schrijfvaardigheid is vakoverschrijdend van tel. Dat geldt ook tijdens het evalueren. Soms kiest men ervoor met de talige kwaliteit van een antwoord geen rekening te houden. Dat geldt bijvoorbeeld wanneer men enkel wil weten of de leerling een rekenvraagstuk kan oplossen. Daartegenover staat dat in het dagelijkse en professionele leven een verzorgde taal steeds van belang is. Het voortdurende gebruik van criteria en rubrieken voor schrijfvaardigheid kan leerlingen hierop attent maken. Daarnaast is het zo dat een essay van een leerling over de rol van water bij het ontstaan van steden in de Middeleeuwen, zowel evaluatiegegevens aanlevert voor het vak Nederlands als voor het vak geschiedenis. Dat betekent dat je twee zaken combineert. Momenteel moeten leerlingen bijvoorbeeld wel rapporten schrijven over de experimenten die ze in natuurkundelessen uitvoeren, maar het schrijven van dergelijke rapporten krijgt zelden aandacht. Het vakoverschrijdende belang van schrijfvaardigheden blijkt uit de voorbeelden in figuur 4.21.

Figuur 4.21: Taken die een beroep doen op schrijfvaardigheden

Natuurwetenschappen. De omgeving waarin levende wezens vertoeven, beïnvloedt hun voorkomen. Levende wezens zijn met andere woorden aangepast aan hun omgeving. Illustreer met een viertal voorbeelden dat dit ook voor jou als mens geldt.

Aardrijkskunde. Op de website van onze gemeente heeft men ruimte voorzien voor informatie voor jongeren. Om de jonge bezoekers met onze gemeente te laten kennismaken wil men een beschrijving geven van de bevolking en de multiculturele samenleving. Hiervoor doet men een beroep op jou. In bijlage vind je een aantal kaarten en grafieken met bevolkinggegevens. Maak op basis daarvan een tekst waarin jij vertelt wat volgens jou belangrijk is om te weten.

Om op deze manier eindtermen aardrijkskunde en natuurwetenschappen te evalueren, moeten leerlingen zelf een antwoord neerschrijven. De kwaliteit van het schrijfproduct moet echter duidelijk onderscheiden worden van de kwaliteit van de andere prestaties. Dat neemt niet weg dat authentieke evaluatietaken beide kunnen combineren. Om die reden willen we een rubriek presenteren die leerlingen van de eerste graad vakoverschrijdend kunnen aanwenden (zie figuur 4.22). Belangrijk is ook dat je als beoordelaar door het gebruik van een dergelijke schrijfrubriek het oordeel over de andere vaardigheden zo weinig mogelijk laat beïnvloeden door de schrijfvaardigheid van de leerling. Om bovenstaande taken te evalueren, zou je dus twee beoordelingsinstrumenten moeten gebruiken, één dat zich richt op de schrijfvaardigheden en één dat de specifieke vaardigheden m.b.t. natuurwetenschappen en aardrijkskunde aan de orde stelt.

Figuur 4.22: Een analytische schrijfrubriek

	5 - goed	4	3 - voldoende	2	1 - onvoldoende
Ideeën en inhoud	De inhoud is duidelijk en gericht. De aandacht van de lezer wordt gevat. Voorbeelden en details verrijken de tekst.		Het onderwerp wordt omschreven maar de uitwerking is nog te algemeen.		Het tekstdoel of thema zijn niet duidelijk. De lezer moet teveel veronderstellen.
Structuur	De structuur ondersteunt de inhoud. De lezer wordt vlot door de tekst geleid.		De lezer wordt zonder teveel afleiding door de tekst geleid.		Er zijn geen duidelijke richting en geen logische achterliggende structuur.
Woordgebruik	Het woordgebruik sluit op een natuurlijke en precieze wijze aan bij de boodschap. Specifieke en accurate woorden maken duidelijk wat de schrijver bedoelt.		De betekenis van wat de schrijver bedoelt wordt goed gecommuniceerd. Het woordgebruik is correct maar brengt het geheel niet op een hoger niveau.		De woordenschat van de schrijver is beperkt. Het woordgebruik draagt niet bij tot de duidelijkheid van het verhaal.

Zinsbouw	De zinnen lezen vlot en zijn goed opgebouwd. De structuur is gevarieerd en creatief.	De zinnen zijn correct. Ze lezen echter eerder mechanisch dan vloeiend.	De zinnen klinken onnatuurlijk en onsamenhangend. De lezer moet teveel moeite doen om het verhaal aan mekaar te knopen.
Conventies	Het correct hanteren van taalconventies verhoogt de leesbaarheid (spelling, hoofdletters, grammatica).	Conventies worden doorgehaars gerespecteerd. Hier en daar verlagen fouten de leesbaarheid.	Herhaaldelijke fouten leiden de lezer af en maken het moeilijk om de tekst te lezen.
Vormgeving	De vorm en presentatie sluiten aan bij de inhoud en dragen bij tot de leesbaarheid van de tekst.	De tekst is verstaanbaar in de huidige vormgeving.	De vormgeving van de tekst schiet tekort en verwtart de lezer.

Om op een authentieke wijze schrijfvaardigheid te evalueren, kan men dus bijvoorbeeld vakoverschrijdend tewerk gaan. Ook binnen het vak Nederlands kunnen echter levensechte taken ontwikkeld worden. Zo verwacht men binnen het vak Nederlands (in de eerste graad van het secundair onderwijs) van leerlingen dat ze kunnen luisteren naar een jeugdprogramma op radio en TV en dat ze een informatief stuk voor bekende leeftijdgenoten kunnen schrijven. Deze luister- en schrijfvaardigheden kunnen aan de hand van de volgende authentieke evaluatietaken beoordeeld worden.

Figuur 4.23: Authentieke evaluatietaken voor luister- en schrijfvaardigheden

Achtergrondinformatie

De mediatheek van onze school heeft een nieuwe documentaire aangekocht. Ook leerlingen zullen die kunnen ontlenu. Jullie weten dat, om op een zinvolle manier gebruik te kunnen maken van de mediatheek, al de documentaires kort in een catalogus zijn samengevat. Voor deze documentaire moet dat nog gebeuren.

De evaluatietaken

Bekijk de documentaire zeer aandachtig. Jullie taak bestaat erin voor deze documentaire een korte beschrijving te maken (ongeveer een halve pagina getypt). Deze samenvatting is bedoeld voor de leerlingen van het eerste en tweede studiejaar. Haal in de samenvatting zeker de volgende elementen aan:

- ✓ Wat is de inhoud van de documentaire? Welke thema's komen aan bod?
- ✓ Wat willen de makers ervan aantonen?
- ✓ Een korte beoordeling: raad je andere leerlingen aan om ze te zien of niet? Waarom (niet)?
- ✓ Gegevens over de documentaire: duur, taal, ondertiteld of niet, wanneer gemaakt en door wie.

Bij het evalueren van de leerlingenprestaties kan een beroep worden gedaan op de volgende taakgebonden analytische rubriek. We onderscheiden hierin drie prestatieniveaus en drie criteria. Om de algemene schrijfvaardigheid van de leerling te evalueren, wordt daarenboven een beroep gedaan op de rubriek bij het vorige voorbeeld (zie fig. 4.24).

Figuur 4.24: Een taakgebonden rubriek

	Excellent	Voldoende	Nog een doel
Samenvatten	De samenvatting dekt de inhoud zeer goed. De lengte is gepast.	De belangrijke aspecten worden in de samenvatting opgenomen. Hoofd- en bijzaken worden voldoende	De samenvatting komt niet goed overeen met de inhoud. Essentiële elementen van de inhoud zijn wegge-

		onderscheiden.	laten.
Begrijpen van de boodschap	De leerling heeft een goed zicht op wat de boodschap of conclusie van de auteur is. Die wordt gevat weergegeven.	De leerling heeft zicht op wat de boodschap of conclusie van de auteur is.	De leerling heeft geen goed zicht op wat de boodschap of conclusie van de auteur is.
Eigen mening weergeven	De beoordeling is relevant en duidelijk.	De beoordeling is duidelijk en relevant maar niet altijd voldoende grondig.	De beoordeling is niet relevant of onduidelijk.

5.2.2. Authentieke evaluatie bij leerlingen in de eerste graad wiskunde

Rekenkundige opgaven worden vaak van hun realistische context ontdaan. Toch is het belangrijk om te weten of leerlingen bepaalde bewerkingen en manieren om problemen aan te pakken ook succesvol kunnen aanwenden in levensechte situaties. Dat geldt onder andere voor de volgende doelstellingen in onze casus: het verband begrijpen tussen de bewerkingen, de volgorde van bewerkingen toepassen, de hoofdbewerkingen correct uitvoeren, en opgaven herformuleren. Om valide evaluatiegegevens te verzamelen en om leerlingen te motiveren, kan daarom een beroep gedaan worden op levensechte evaluatietaken zoals in figuur 4.25.

Figuur 4.25: Een authentieke rekentaak

Achtergrondinformatie
Zaterdagavond organiseer ik (de leerkracht) een tuinfeestje. Gezellig met vrienden wat eten en drinken. Dingen om te drinken heb ik nog voldoende in huis. Wat het eten betreft, moet ik alles nog aankopen en daarom zit ik een beetje met de handen in het haar. Zelf had ik gedacht aan een bescheiden barbecue. Ik heb een budget van 50€ en weet niet wat ik hiervoor kan kopen. Mag ik jou vragen mij even te helpen rekenen?

De evaluatietask
Stel dat jij een van de gasten bent, wat zou je mij dan adviseren om aan te kopen? Hiertoe moet je een lijst opstellen met inkopen die ik met dat bedrag kan doen. Dit moet je ook nog weten. Er komen 5 gasten. Met mezelf er bij moet ik dus voor 6 personen eten voorzien. Voor iedereen zou er wat vlees en wat groenten voorzien moeten worden. Ik kan echter alleen kopen wat er de reclamefolder staat die jullie hebben gekregen en ik mag maximaal 50 € uitgeven.

Stel een winkellijst op. Let hierbij op het volgende:

- ✓ Geef goed aan hoeveel je van welk product koopt en hoeveel elke aankoop apart kost.
- ✓ Gebruik minstens twee van de kortingsbonnen uit de folder.
- ✓ Geef duidelijk aan wat de totale prijs en de prijs per persoon zal zijn.
- ✓ Noteer op een duidelijke manier alle rekenbewerkingen die je doet.
- ✓ Kies zelf wat je lekker vindt.

Om een taak zoals deze op te lossen, moeten leerlingen twee te onderscheiden groepen van vaardigheden beheersen. Ze moeten het geformuleerde probleem kunnen aanpakken (begrijpen, plannen, oplossen en beantwoorden) en ze moeten de berekeningen op een correcte manier kun-

nen uitvoeren. Om de eerste groep vaardigheden te beoordelen, presenteren we een rubriek die bij de aanpak van alle rekenproblemen kan worden gebruikt. De specifieke doelstellingen die we met deze taak willen evalueren, beoordelen we aan de hand van een holistische rubriek in figuur 4.26.

De doelstellingen die we met de authentieke rekentaak willen evalueren, vormen slechts één aspect van de rubriek in figuur 4.26. Ze richten zich meer bepaald op het correct oplossen van het probleem. In principe kan elk van de criteria verder gespecificeerd worden. Met deze taak willen we leerlingen echter beoordelen op hun inzicht in de bewerkingen en de mate waarin ze de hoofdbewerkingen correct kunnen uitvoeren. Gesteld dat we hieromtrent een algemeen waardeoordeel willen uitspreken, dan kan dat aan de hand van de holistische rubriek die we in figuur 4.11 presenteerden. Om evaluatiegegevens te verzamelen, kunnen we een beroep doen op de bijhorende evaluatielijst.

Figuur 4.26: Rubriek voor het aanpakken van rekenproblemen

	3 - goed	2 - voldoende	1 - onvoldoende
Verstaan van het probleem	Al de belangrijke elementen in het probleem worden geïdentificeerd. Hoofd- en bijzaken worden onderscheiden.	Het probleem wordt grotendeels begrepen. Een aantal relevante elementen wordt echter over het hoofd gezien.	Het probleem is niet begrepen.
Plannen van de oplossing	Het plan heeft alles in zich om op een efficiënte manier tot een correcte oplossing te komen.	Het plan kan tot de correcte oplossing leiden maar is niet zo efficiënt.	Het plan is niet geschikt om tot een correcte oplossing van het probleem te komen.
Oplossen van het probleem	De oplossing getuigt van een goed inzicht in de nodige concepten en van een perfect uitvoeren van de berekeningen.	De oplossing getuigt van voldoende inzicht in de nodige concepten en van correct uitvoeren van de berekeningen. Toevallige misrekeningen kunnen voorkomen.	De leerling beheerst de noodzakelijke rekenvaardigheden onvoldoende om tot de juiste oplossingen te komen. Er worden teveel fouten gemaakt.
Beantwoorden van het probleem	Het antwoord is volledig met een duidelijke, coherente verklaring.	Het antwoord volstaat maar de uitleg is deels onduidelijk of sluit niet volledig aan bij de vraag.	Het antwoord is onduidelijk of vormt geen antwoord op het probleem.

5.2.3. Authentieke evaluatie bij leerlingen Elektriciteit/Elektronica

Uit het geheel van doelstellingen voor de leerlingen Elektriciteit/Elektronica selecteren we een drietal waarvoor we een geschikte vaardigheidsevaluatie zullen ontwikkelen, namelijk: de opbouw van een stof kunnen toelichten, meettoestellen (volt-, ampère- en ohmmeter) kunnen gebruiken en een verslag kunnen maken van de metingen die zij zelf uitvoeren.

Figuur 4.27: Een authentieke elektriciteits oefening

Achtergrondinformatie

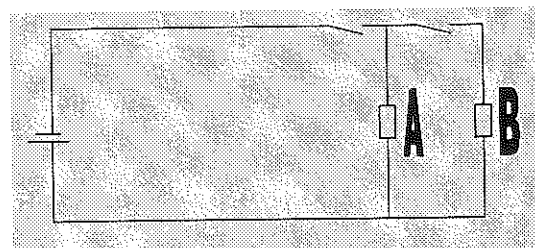
De leerlingen krijgen een praktische opstelling van een elektrische stroomkring als uitgangspunt voor deze evaluatietask. (Eventueel stellen ze de kring eerst zelf op). Alle gegevens over de

stroombron zijn gekend. Enkel de gegevens van de twee weerstanden zijn onbekend. In feite bestaan de twee weerstanden uit verschillende stoffen maar dat is van buitenaf niet te zien, de ene weerstand (A) bestaat gewoon uit koperdraad terwijl de andere weerstand (B) effectief een eigenlijke weerstand inhoudt. De leerlingen hebben zowel een volt-, ampère- als een ohmmeter ter beschikking.

De evaluatietask

Je bent mede verantwoordelijk voor het onderzoek naar geleiders in de fabriek van de gebroeders Antheunis. Het doel van jullie onderzoek is de ontdekking van een supergeleider (dat is een geleider met een minimale weerstand). Het beste resultaat dat jullie tot nu toe geboekt hebben, is verwerkt in weerstand A.

Vanmorgen is een onbekende man naar jullie fabriek gekomen. De man beweert dat hij een geleider (B) heeft ontdekt die nog minder elektrische weerstand heeft dan wat jullie tot nog toe gerealiseerd hebben (A dus). Jouw baas vraagt je om na te gaan of het verhaal van de man klopt. Hij vraagt je ook een verslag te schrijven van je bevindingen. Jouw opdrachten zijn m.a.w.:



1. Toon met het gebruik van de juiste meettoestellen aan welke van beide weerstanden de minste elektrische weerstand kent. Geef op de tekening aan hoe je de meettoestellen hebt aangesloten.
2. Maak een verslag van alle stappen en metingen die je doet.
3. Los van de metingen die je doet, vraagt de baas je ook om hem te verklaren waarom het ene materiaal de stroom beter geleidt dan de andere. Leg dit aan de hand van jouw kennis van de opbouw van een stof uit.

5.2.4. Authentieke evaluatie bij leerlingen Verzorging

Studierichtingen die praktijkoefeningen en stages inrichten, hebben het voordeel dat er levens-echte situaties en contexten gecreëerd worden. Op die manier wordt het mogelijk op een authentieke manier de kennis en vaardigheden van leerlingen te benaderen. Dat kan aan de hand van gecreëerde producten zoals bij de leerlingen Elektriciteit/Elektronica of Houtbewerking maar ook door het gedrag van leerlingen rechtstreeks te observeren. Bij de leerlingen Verzorging gebeurt dat bijvoorbeeld op volgende manier.

Een leerling die stage loopt, ondergaat veel levensechte vaardigheidstaken. De leerling draait mee in de dagelijkse realiteit waarin continue kennis moet worden aangewend en vaardigheden gedemonstreerd. Leerlingen die stage lopen in een kleuterschool doen een beroep op hun sociale vaardigheden, technische vaardigheden, persoonlijkheidsaspecten en attitudes. De stage zelf is dus een authentieke evaluatietask. Observatie van de leerling tijdens deze momenten kan een schat aan evaluatiegegevens opleveren, ten minste wanneer de observatie gestructureerd wordt. Leerlingen en beoordelaars moeten weten wat er verwacht wordt en wat geëvalueerd zal wor-

den. Alvorens algemene beoordelingen uit te spreken, worden daartoe gegevens verzameld aan de hand van checklists en evaluatielijsten. We presenteren eerst een holistische rubriek en daarna een beoordelingsinstrument dat het mogelijk maakt gericht evaluatiegegevens te verzamelen.

Figuur 4.28: Een holistische rubriek voor stage-evaluatie in de kleuterschool

Uitstekend. De leerling geeft blijk van een prima beheersing van al de sociale vaardigheden. Dit zowel ten aanzien van de kleuters als van collega kleuterleidsters. De leerling is trots en bekommerd om de kwaliteit van de stageactiviteiten. Er is sprake van een zeer sterke beroepsmotivatie. Het voorkomen, de werkhouding en de communicatievaardigheden zijn uitstekend. Technische vaardigheden worden zeer goed beheerst en gepast aangewend. Pedagogische taken worden op een zeer goede manier uitgevoerd.

Goed. De leerling beheerst de sociale vaardigheden. De leerling is trots en bekommerd om de kwaliteit van de stageactiviteiten. Het voorkomen, de werkhouding en de communicatievaardigheden zijn goed. Technische vaardigheden worden beheerst en gepast aangewend. Pedagogische taken worden goed uitgevoerd.

Voldoende met leemtes. De leerling beheerst de belangrijkste sociale vaardigheden en geeft doorgaans blijk van voldoende motivatie. Het voorkomen, de werkhouding en de communicatievaardigheden zijn voldoende maar een aantal aspecten kunnen beter. De meeste technische vaardigheden worden voldoende beheerst. Het gros van de pedagogische taken wordt goed uitgevoerd.

Onvoldoende. Belangrijke sociale vaardigheden worden niet beheerst. De beroepsmotivatie ontbreekt op belangrijke momenten. Het voorkomen, de werkhouding en de communicatievaardigheden moeten duidelijk beter. De meeste technische vaardigheden worden onvoldoende beheerst. Pedagogische taken worden doorgaans onvoldoende goed uitgevoerd.

Alvorens met deze rubriek beoordelingen uit te spreken over de prestaties van leerlingen, moet op een analytische manier gegevens verzameld worden. Dit kan aan de hand van de evaluatielijst in figuur 4.29 (gebaseerd op SLI, 2002). De lijst geeft (een selectie van) criteria waaraan een goede stage moet voldoen. Elk van deze criteria wordt beoordeeld met de volgende descriptieve schaal: uitstekend, goed, voldoende met leemtes en onvoldoende. Beoordelaars krijgen hierbij voldoende mogelijkheden om opmerkingen te formuleren.

Figuur 4.29: Evaluatielijst voor stage-evaluatie in de kleuterschool

Sociale vaardigheden	
✓	Neemt spontaan contact op met de kleuters.
✓	Neemt een correcte houding aan (respectvol, vriendelijk, hulpvaardig).
✓	Stimuleert en motiveert tot zelfredzaamheid.
✓	Legt goede contacten met collega's.
✓	Kan met collega's opdrachten evalueren.
✓	Stelt vragen.
✓	Geeft informatie door.
✓	Is in staat tot aangename samenwerking.
Beroepsmotivatie	
✓	Is bekommerd om het goede verloop van de stage.

- ✓ Doet het werk graag.
- ✓ Kan fouten aanvaarden.

Houding, voorkomen en communicatievaardigheden

- ✓ Heeft een verzorgd voorkomen en heeft daar ook aandacht voor tijdens het werken.
- ✓ Kan zich uitdrukken in een behoorlijke taal.
- ✓ Luistert aandachtig naar opdrachten.

Werkhouding

- ✓ Heeft verantwoordelijkheidszin.
- ✓ Is stipt.
- ✓ Komt op voor de eigen mening.
- ✓ Neemt initiatief.

Technische vaardigheden

- ✓ Past handhygiëne toe.
- ✓ Verzamelt informatie over de kleuter.
- ✓ Kan instaan voor hygiënische zorgen (handjes wassen, verschonen, hulp bij toilet, ...).
- ✓ Helpen bij het eten.

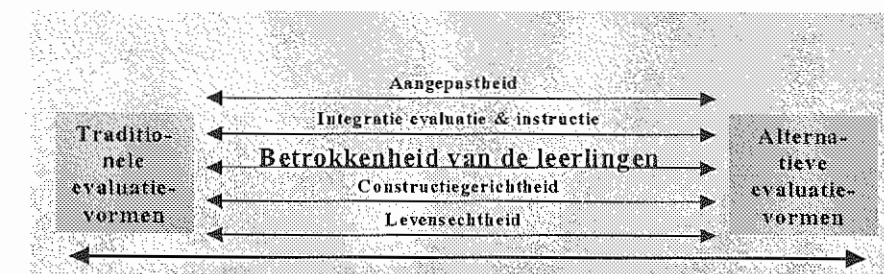
Pedagogische taken

- ✓ Toezicht en begeleiding – heeft aandacht voor:
 - Ruzies (voorkomen, oplossen)
 - Gevaarlijke handelingen
 - Observeren en stimuleren van spel
- ✓ Kinderen troosten, knuffelen, opbeuren.
- ✓ Stimuleert het aanleren van goede gewoontes.
- ✓ Voert creatieve opdrachten uit (vertellen, zingen, knutselen).

Hoofdstuk 3: De betrokkenheid van leerlingen

In de toetscultuur ligt de verantwoordelijkheid voor het evaluatieproces bij de leerkracht. De veranderende visies op evaluatie en de verruimde functies van evaluatie maken echter dat alternatieve evaluatievormen - waarin de rol van de leerlingen sterker benadrukt wordt - opgang maken. Bij het beschrijven van de kanteling van een traditionele toetscultuur naar een assessmentcultuur werd de verantwoordelijkheid van leerlingen tijdens de evaluatie als een belangrijke evolutie naar voren geschoven. In dit hoofdstuk staat daarom de vraag wie welke rol kan vervullen tijdens het evaluatieproces centraal. Hierbij denken we aan *zelfevaluatie*, *peerevaluatie* en *collaboratieve evaluatie*. Aan elk van deze evaluatievormen besteden we in dit hoofdstuk aandacht.

Zelfevaluatie, peerevaluatie en collaboratieve evaluatie zijn vormen van alternatieve evaluatie in die zin dat ze tegemoetkomen aan één van de eisen die geformuleerd werden ten aanzien van alternatieve evaluatievormen, namelijk de betrokkenheid van de leerlingen.



Figuur 4.30: Leerlingenbetrokkenheid als kenmerk van alternatieve evaluatievormen

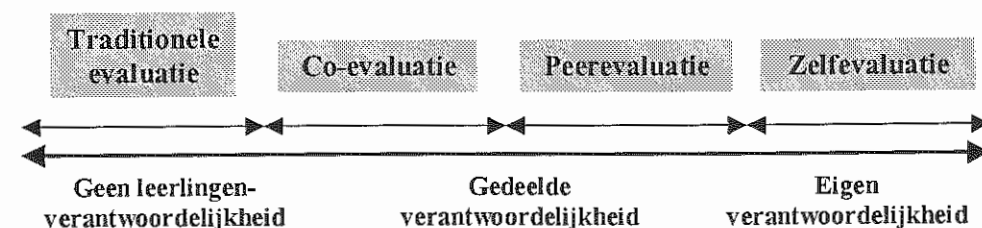
Voor de betrokkenheid van leerlingen geldt hetzelfde als voor andere kenmerken van alternatieve evaluatievormen: een grote betrokkenheid van leerlingen is een voldoende maar geen noodzakelijke voorwaarde om van een alternatieve evaluatie te kunnen spreken. De andere kenmerken van alternatieve evaluatievormen worden daartoe buiten beschouwing gelaten wanneer we het hebben over de betrokkenheid van leerlingen. Dat betekent dat vormen van zelf- en peerevaluatie op de andere kenmerken van alternatieve evaluatievormen zeer traditioneel kunnen scoren. Men kan leerlingen bijvoorbeeld zelfstandig hun eigen antwoorden (of deze van klasgenoten) op een reproductieve invuloefening laten evalueren. Leerlingen worden op die manier wel betrokken bij de evaluatie maar niet op een manier zoals de andere kenmerken van alternatieve evaluatie dat vereisen. Of leerlingen al dan niet betrokken worden in de evaluatie, impliceert dus niet dat automatisch aan andere kenmerken tegemoet wordt gekomen. Uit onderstaande beschrijving zal bovendien blijken dat de betrokkenheid van leerlingen -net als de andere

kenmerken van alternatieve evaluatievormen- kan worden opgevat als een continuüm en op verschillende manieren kan worden ingevuld.

De mate waarin leerlingen betrokken zijn bij en verantwoordelijk zijn voor de evaluatie kan sterk variëren. Dit geldt voor elke fase van het evaluatieproces: het vooropstellen van doelstellingen, het ontwikkelen of selecteren van criteria, het gebruiken van de criteria en het voorzien van feedback.

De verantwoordelijkheid voor de evaluatie kan of volledig bij de leerkracht of volledig bij de leerlingen liggen. Meestal zal er evenwel sprake zijn van een gedeelde verantwoordelijkheid. Schematisch kunnen we dat als volgt voorstellen (Harris & Bell, 1994).

Figuur 4.31: Een continuüm van leerlingenbetrokkenheid



Men zou zelfevaluatie kunnen opvatten als het eindpunt van een proces waarin leerlingen steeds minder gestuurd en gereguleerd worden bij de evaluatie. De leerkracht leert de leerlingen geleidelijk aan om hun eigen prestatie af te wegen tegen vooropgestelde criteria en zelf op zoek te gaan naar hiaten in het leerproces. Het eindstadium wordt bereikt wanneer leerlingen volledig de verantwoordelijkheid dragen over hun leer- en evaluatieproces. Van dan af worden leerlingen gestimuleerd om zelf acties te ondernemen en hun eigen leerproces bij te sturen. In het schema wordt tevens duidelijk dat ook binnen de onderscheiden evaluatievormen nog aanzienlijke verschillen kunnen bestaan wat de mate van leerlingenverantwoordelijkheid betreft. Hetzelfde geldt voor de invulling van deze verantwoordelijkheden. In de beschrijving van zelfevaluatie, peerevaluatie en collaboratieve evaluatie zal dit duidelijk worden.

1 Verschillende vormen van leerlingenbetrokkenheid

Leren veronderstelt dat de lerende wordt aanzet om actief nieuwe informatie te manipuleren. Enkel het opdoen en gebruiken van nieuwe ideeën in verscheidene contexten is echter niet voldoende om te garanderen dat effectief en langdurig leren plaats vindt (Bartels, 1998). Als men leerlingen ertoe wil brengen hun eigen leren in handen te nemen, dan moeten manieren gevonden worden om de bewustwording van dat leren en de verantwoordelijkheid ervoor te ontwikkelen. Het doelbewust opnemen van leerlingenbetrokkenheid in de leerprocessen kan deze doelen helpen realiseren. In de volgende paragrafen geven we eerst een illustratie van leerlingenbe-

trokkenheid. Daarna definiëren we een aantal specifieke evaluatievormen die de leerling activeren en medeverantwoordelijk maken voor het evaluatieproces.

1.1 De betrokkenheid van leerlingen geïllustreerd

Om te illustreren hoe leerlingen betrokken kunnen worden bij het beoordelen van hun eigen prestaties, richten we ons op het vak wiskunde. We spitsen ons toe op een specifiek aspect van het wiskundecurriculum, namelijk het maken, interpreteren en gebruiken van grafieken. In feite kunnen voor tientallen andere taken gelijkaardige zelfevaluatieinstrumenten ontwikkeld worden. Te denken valt aan het oplossen van problemen, werken in groep, verhalend schrijven, plannen van activiteiten etc. Voor deze taken kan telkens een beoordelingsinstrument samengesteld worden dat door leerlingen zelf gebruikt kan worden. In het voorbeeld van het maken van grafieken gebeurt dit in de vorm van een evaluatielijst (Hibbard, 1996). Die wordt door de leerkracht vóór de taak toegelicht om aan de leerlingen duidelijk te maken wat het relatieve belang is van de verschillende leeruitkomsten en criteria. Vervolgens wordt de evaluatielijst aan de leerlingen bezorgd. Na het maken van de taak wordt de leerlingen gevraagd om hun eigen prestaties of producten (de gemaakte grafiek) te beoordelen aan de hand van de evaluatielijst. De leerkracht evalueert vervolgens zowel de taken als de zelfevaluatie van de leerlingen.

Het is mogelijk om met betrekking tot een aantal kenmerken van de doelgroep aangepaste evaluatieinstrumenten te ontwikkelen. We illustreren dit kort met de volgende twee instrumenten.

- ✓ In het lager onderwijs onderscheidt men drie prestatieniveaus, namelijk P (prachtig), O (OK) en N (nog oefenen). Voor elk van de in de taak onderscheiden dimensies wordt dit onderscheid in prestatieniveaus gemaakt, bijvoorbeeld:

Lijnen of balken tekenen op de grafiek

- ✓ P: De getekende lijnen of balken zijn nauwkeurig en zeer net.
- ✓ O: Er zijn enkele fouten of het werk ziet er wat slordig uit.
- ✓ N: Er zijn veel fouten en het werk is zeer slordig.

Labelen van de data

- ✓ P: Elke as heeft een naam en is onderverdeeld in eenheden.
- ✓ O: De as heeft duidelijkere namen nodig en/of moet in eenheden worden onderverdeeld.
- ✓ N: De namen van de assen en de onderverdeling in eenheden ontbreken.

De leerling kan aan de hand van de schaal uitspraken doen over de eigen grafiek of over de grafiek van een klasgenoot. Een beoordeling op de volledige lijst van criteria maakt het mogelijk om een duidelijk onderscheid te maken tussen goede en minder goede grafieken en om aan te tonen wat de sterke en zwakke punten zijn van een specifieke grafiek.

- ✓ In het secundair onderwijs gebeurt de zelfevaluatie op basis van een evaluatielijst met een differentieerbare schaal. Leerlingen beoordelen de zelf gemaakte grafieken of deze van medeleerlingen door waarderingen uit te spreken over de verschillende criteria.

Figuur 4.32: Voorbeeld van een zelfevaluatie-instrument

	mogelijke evaluatiepunten	zelfevaluatie
De titel van de grafiek geeft duidelijk het verband aan tussen de assen.	_____	_____
Er werd een geschikt grafiektype gebruikt.	_____	_____
De assen dragen een duidelijk label.	_____	_____
De lijnen of balken gebruiken de beschikbare ruimte goed.	_____	_____
De grafiek is makkelijk te interpreteren	_____	_____
totaal	_____	_____

In de verschillende jaren van het voortgezet onderwijs wordt een combinatie gebruikt van een rubriek en aan het niveau van de leerlingen aangepaste evaluatielijsten. De criteria stellen de leerlingen in staat de kwaliteit van de afzonderlijke taken te beoordelen terwijl de rubriek een algemeen beeld geeft van de kwaliteit van de grafieken die de leerling ontwerpt. De rubriek die aansluit bij bovenstaande evaluatielijsten hebben we reeds gepresenteerd in figuur 3.9.

Het gebruik van deze vooraf ter beschikking gestelde evaluatieinstrumenten wordt naar het einde van het secundair onderwijs vaak gradueel afgebouwd. De leerlingen gebruiken dan niet meer de evaluatielijsten van de leerkracht maar leren de door henzelf samengestelde evaluatielijsten hanteren.

1.2 Vormen van leerlingenbetrokkenheid gedefinieerd

Bij de traditionele evaluatiemethoden is het de leerkracht die verantwoordelijk is voor wat er geleerd moet worden, welke criteria bij de evaluatie gebruikt worden en de eigenlijke evaluatie. Vanuit het besef dat leerlingen niet alleen moeten leren maar ook moeten weten hoe te leren en hoe te reflecteren over dat leren, evolueert het evaluatieproces naar een alternatief model. Dat model omvat een herverdeling van de verantwoordelijkheden ten aanzien van de onderwijsdoelen, het curriculum en de evaluatie. Omwille van de gewichtige beslissingen die vaak op basis van een evaluatie genomen worden, is inspraak van de leerlingen en samenwerking net bij de evaluatie het meest belangrijk. Verschillende evaluatievormen komen hieraan tegemoet. We maken een onderscheid tussen zelfevaluatie, peerevaluatie en collaboratieve evaluatie. De betrokkenheid en de rol van de leerlingen in de evaluatie is telkens anders. Dat blijkt uit de volgende definities.

Zelfevaluatie verwijst naar het evalueren door leerlingen van hun eigen kennis, vaardigheden, producten, leerprocessen en attitudes. De leerling richt zich systematisch op

de eigen prestaties, meestal met als doel de toekomstige prestaties te verbeteren. De leerling krijgt hiertoe criteria, rubrieken en standaarden ter beschikking of ontwikkelt die zelf. Zelfevaluatie kan ook bestaan uit het bekritisieren of simpelweg beschrijven van de eigen prestaties (ASCD, 1996; Rolheiser et al., 2000; Mabry, 1999).

Peerevaluatie. Een 'peer' is gewoonlijk een leeftijdsgenoot of iemand die in hetzelfde leerjaar zit. Leerlingen in dezelfde klas zijn dus mekaars peers. Peerevaluatie kan dan ook eenvoudig gedefinieerd worden als het evalueren door leerlingen (peers) van elkaars werk (Tondeur, 1998).

Collaboratieve evaluatie (of co-evaluatie) is een tussenstadium tussen de traditionele eenzijdige evaluatie door de leerkracht enerzijds en zelf- en peerevaluatie anderzijds. In een collaboratieve evaluatiemethode komen de leerling, de medeleerlingen en de leerkracht samen om te evalueren. Collaboratieve evaluatie creëert aldus een dialoog tussen de leerkracht en de leerlingen tijdens het beoordelen. De leerling is niet de eindverantwoordelijke voor de evaluatie maar collaboreert wel in het evaluatieproces (Harris & Bell, 1994).

Het onderscheid tussen de verschillende vormen van leerlingenbetrokkenheid situeren wij in de eigenlijke beoordelingsfase. Het gaat erom wie instaat voor de beoordeling. Andere benaderingen doen dat niet. De term zelfevaluatie wordt door sommigen enkel weerhouden voor evaluatiecontexten waarin leerlingen hun eigen leerdoelen bepalen, zelf criteria ontwikkelen en hun eigen prestaties beoordelen. In onze benadering maakt het niet uit waar de doelstellingen vandaan komen en wie de evaluatiecriteria en -instrumenten ontwikkelde. Pas wanneer de criteria en instrumenten gebruikt worden -bij het beoordelen dus- wordt het onderscheid tussen de verschillende vormen van leerlingenbetrokkenheid duidelijk. De vraag is dan door wie de leerling beoordeeld wordt: (1) door zichzelf, (2) door peers, of (3) gezamenlijk door de leerlingen zelf, de klasgenoten, de leerkracht en eventuele derden.

1.3 Kenmerken van de verschillende vormen van leerlingenbetrokkenheid

Uit de definities blijkt dat het betrekken van leerlingen in de evaluatie verschillende vormen kan aannemen. Elke vorm heeft een aantal specifieke kenmerken.

1. Zelfevaluatie

Alle vormen van zelfevaluatie hebben een gemeenschappelijk aspect: ze vragen leerlingen om hun eigen werk te beoordelen. Men vraagt leerlingen om te reflecteren over hun eigen werk en om te bepalen wat ze geleerd hebben, welke onderdelen nog onduidelijk zijn en welke aspecten nog bijkomende aandacht verdienen. Hiertoe wordt leerlingen de tijd gegeven om aandachtig hun eigen leerproces en vooruitgang te evalueren. Bij een zelfevaluatie hebben de leerlingen een

organisatorische, sturende en inhoudelijke rol. Dat kan bijvoorbeeld inhouden dat de leerlingen zelf bepalen hoe, wanneer en op basis waarvan ze beoordeeld zullen worden.

De manier waarop leerlingen uitgenodigd worden om hun eigen prestaties te evalueren, kan verscheidene vormen aannemen:

- ✓ schriftelijke besprekingen;
- ✓ discussie (in groep of met de hele klas);
- ✓ logboeken;
- ✓ checklists en vragenlijsten;
- ✓ leerkracht-leerlingen interviews, onderlinge leerlingeninterviews.

Zelfevaluatie als evaluatiemethode is veelal formatief van karakter. Niet enkel het beoordelen is immers belangrijk, ook reflectie speelt een belangrijke rol. Van dit reflectieproces verwacht men dat het leidt tot verandering in het denken, de houding en het gedrag van leerlingen (Dochy & Segers, 1999; Rolheiser, 2000).

2. Peerevaluatie

Terwijl zelfevaluatie van leerlingen verwacht dat ze hun eigen werk bestuderen, is het bij peerevaluatie zo dat de leerlingen zich richten op het werk van anderen. Kenmerkend voor peerevaluatie is dat elk lid van de (klas)groep kritisch kijkt naar de activiteiten van de anderen, over deze prestaties discussieert en ze evalueert. Peerevaluatie kan op verschillende manieren gestructureerd worden en het uitzicht van de eigenlijke evaluatie kan dan ook sterk variëren. In principe is het zo dat methoden en instrumenten die voor zelfevaluatie gebruikt worden, eveneens geschikt zijn om de prestaties van anderen te evalueren. Een vaak gehanteerde werkwijze bestaat uit het invullen van een evaluatielijst die een aantal vragen (criteria) over de te evalueren leerlingprestaties aan de orde stellen.

3. Collaboratieve evaluatie

Bij collaboratieve evaluatie bepaalt de leerkracht samen met de leerlingen de doelen van het leerproces, de manier waarop die doelen bereikt moeten worden en de wijze waarop ze geëvalueerd en beoordeeld worden. Leerlingen worden dus in het evaluatieproces betrokken maar het blijft de leerkracht die stuurt en de eindverantwoordelijkheid heeft (Harris & Bell, 1994; Topping, 1998).

In paragraaf 3 van dit deel bespreken we het verloop van een evaluatieproces dat de betrokkenheid van leerlingen zoveel mogelijk wil aanhalen. Om de kenmerken van collaboratieve evaluatie te schetsen, is het echter leerrijk hier reeds een blik te werpen op de verschillende stappen waaruit een collaboratief evaluatieproces bestaat. In het verloop van een collaboratieve evaluatie kunnen een zestal stappen onderscheiden worden. Deze zijn:

1. komen tot en verduidelijken van evaluatiecriteria;

2. bereiken van overeenkomst over de puntenverdeling (wie gaat welk aandeel van de punten verdelen?);
3. zelf- en/of peerevaluatie door de leerlingen;
4. evaluatie van de leerlingen door de leerkracht;
5. vergelijken, contrasteren en bediscussiëren van beider evaluatiegegevens;
6. komen tot evaluatieve uitspraken en feedback.

Het feit dat er sprake is van verschillende stappen hoeft niet te impliceren dat deze ook chronologisch moeten worden doorlopen. Essentieel is wel dat in elke vorm van collaboratieve evaluatie wordt besproken wie evalueert (peers, leerling, leerkracht), welke de proces- en productcriteria zijn en op welke manier de puntenverdeling gebeurt. Verschillen hieromtrent leiden tot verscheidene modellen en methoden van collaboratieve evaluatie. Die komen later in dit hoofdstuk aan bod.

2. Waarom leerlingbetrokkenheid?

Er kan een verscheidenheid aan rechtvaardigingen van en stimulansen tot het betrekken van leerlingen worden gevonden in de evaluatieliteratuur. Vele elementen hebben onmiddellijk betrekking op pedagogische aspecten van de hedendaagse onderwijsbenadering. Te denken valt aan levenslang leren, het ontwikkelen van metacognitie, de verantwoordelijkheid van leerlingen voor hun eigen leerproces en -product en dergelijke. Doorheen hun professionele leven zullen leerlingen immers genoodzaakt zijn de kwaliteit van het werk van ondergeschikten, collega's en uiteraard ook van zichzelf op een kwaliteitsvolle manier te beoordelen. In een eerste paragraaf rapporteren we een aantal argumenten die kaderen in dit pleidooi voor leerlingbetrokkenheid. Daarna wijzen we erop dat toch ook enige terughoudendheid aan de orde is.

2.1 Een pleidooi voor leerlingenbetrokkenheid

Een kwaliteit die van grote invloed is op duurzaam leren, is kunnen reflecteren op het eigen gedrag. *Leren reflecteren* is daarom de belangrijkste motivatie om leerlingen verantwoordelijkheid te geven in hun eigen leer- en evaluatieproces. Daarnaast zijn er nog aantal andere *onderwijskundige én praktische voordelen* van leerlingenbetrokkenheid.

2.1.1. Leren reflecteren

Om van bepaalde ervaringen te komen tot leren, is het noodzakelijk dat de leerlingen reflecteren. Reflectie kunnen we definiëren als 'ideeën en conclusies die het resultaat zijn van het nadenken over het eigen werk'. Die ideeën zijn gebaseerd op expliciete criteria en kunnen helpen toekomstige doelen en acties te plannen en te ondernemen. We spreken van reflectie wanneer leerlingen nagaan in welke mate hun prestaties tegemoet komen aan vooropgestelde criteria en

wanneer ze de effectiviteit van hun inspanningen beoordelen en plannen maken voor verbetering. Een externe evaluatie door een expert is in dat opzicht minder waardevol dan het eigen aanvoelen van de leerlingen of iets al dan niet beheerst wordt (Bartels, 1998; Rolheiser et al, 2000).

Om zich te ontwikkelen tot een autonome lerende moet de leerling geleidelijk aan strategieën uitproberen die het mogelijk maken om zowel afstand te nemen van de eigen perspectieven en werkwijze als om de eigen sterktes en zwaktes te herkennen. Men moet leerlingen leren hoe ze zichzelf kunnen observeren, hoe ze hun eigen gedrag en prestaties kunnen afwegen tegen bepaalde standaarden en hoe ze kunnen zoeken naar geschikte methodes om zichzelf indien nodig te corrigeren (Bartels, 1998; Rolheiser et al, 2000). Reflecteren op het eigen leren is immers een krachtig instrument om nieuwe kennis en vaardigheden te integreren in de bestaande denk- en handelingsschema's (Clemens et al, 1999). Het voornaamste doel van het betrekken van leerlingen in het evaluatieproces is het ontwikkelen van het vermogen om het eigen leerproces en -product te beoordelen (Conley, 1996). Vertrouwd worden met het eigenlijke evaluatieproces is dan ook even belangrijk als de conclusies waartoe de evaluatie leidt (Zariski, 1996).

2.1.2. Positieve onderwijskundige gevolgen van leerlingenbetrokkenheid

Een grote betrokkenheid van de leerlingen bij het eigenlijke evaluatieproces heeft nog een aantal andere nastrevenswaardige doelen. Het belang van die gevolgen wordt ondersteund door een aantal praktijkervaringen die stellen dat er een directe correlatie is tussen de mate waarin leerlingen evaluatievaardigheden onder de knie hebben en de kwaliteit van het werk dat ze realiseren (Zariski, 1996; de Rijke & Beemer, 1995; Mabry, 1999). Er zijn met andere woorden aanwijzingen voor het feit dat die vaardigheden leiden tot een gewichtige verbetering van de prestaties van leerlingen. De volgende elementen spelen hierin een rol.

Leerlingen gedragen zich als quasi-professionelen en gaan zichzelf zo ook ervaren. Ze denken na over wat belangrijk is om van kwaliteitsvolle prestaties te kunnen spreken en ontwikkelen hun eigen kwaliteitsstandaarden. Ze ontwikkelen met andere woorden vaardigheden en attitudes die van tel zijn in hun dagelijks leven, hun studieloopbaan en hun latere beroepsuitoefening. Betrokkenheid en het ervaren van verantwoordelijkheid voor de eigen ontwikkeling versterken bovendien de motivatie. De Rijke & Beemer (1995) schrijven dat alleen al het feit dat men verantwoordelijkheid krijgt en niet alleen een evaluatie ondergaat, motiverend werkt en energie geeft. De leerlingen krijgen tijdens de zelfevaluatie het gevoel dat ze invloed en zelfs een beetje macht hebben. Deze motivatie maakt dat leerlingen gaan investeren in en verantwoordelijkheid opnemen voor hun eigen leren. Op die manier kan soms op een spectaculaire manier het eigen kunnen verbeterd worden. De participatie van leerlingen aan het evaluatieproces en de daaruit resulterende betrokkenheid maken verder dat leerlingen gemakkelijker gebruik gaan maken van diepgaande verwerkingsstrategieën en metacognitieve strategieën (Topping, 1998).

Specifiek bij peerevaluatie kunnen leerlingen bovendien ideeën opdoen door elkaars werk te bestuderen. Bovendien zou peerevaluatie tevens hun ondernemingszin, motivatie en respect voor mekaar verhogen (Mabry, 1999). Kortom: de betrokkenheid van leerlingen maakt dat het evaluatiemoment ook een leermoment wordt. Het leren gebeurt dus al tijdens de eigenlijke evaluatie en niet alleen na de terugkoppeling van een beoordeling. Instructie en evaluatie verlopen geïntegreerd.

2.1.3. Praktisch voordelen

Het laatste element dat we willen aanhalen, is zowel een argument voor zelfevaluatie, peerevaluatie als collaboratieve evaluatie. Evaluatiemethoden zoals vaardigheidsevaluatie en authentieke evaluatie bieden een aantal speciale voordelen, maar ze zijn tijdrovend. De voordelen kunnen echter nog vergroot worden en de tijdsinvestering kan meer behapbaar worden indien de verantwoordelijkheden van de evaluatie gedeeld worden (Mabry, 1999; Zariski, 1996). De meest geschikte manier om verantwoordelijkheden te delen bestaat uit het activeren van leerlingen. Leerlingen kunnen individueel hun eigen prestaties beoordelen of die van hun klasgenoten. Uiteraard kan ook in groep worden samengewerkt om het eigen werk en dat van klasgenoten te evalueren. Op die manier kan de werklast van de leerkracht verminderd worden. Naast de pedagogische voordelen van het betrekken van leerlingen in de evaluatie zijn er dus eveneens praktische voordelen. De vragen of, wanneer, hoe en met wie evaluatieverantwoordelijkheden best gedeeld worden zijn echter afhankelijk van de specifieke evaluatiecontext. Ze lenen zich dus niet tot uniforme antwoorden.

Een ander praktisch voordeel van leerlingenbetrokkenheid is het feit dat de rechtvaardigheid van de evaluatie mede door de leerlingen zelf bewaakt kan worden. De acceptatie en de legitimiteit van het eindoordeel van de evaluatie is wellicht groter naarmate de betrokkene actiever heeft bijgedragen aan het verloop van de evaluatie en de eigenlijke beoordeling.

2.2 Toch ook enige terughoudendheid

Naast bovenstaande argumenten die pleiten voor leerlingenbetrokkenheid kunnen ook een aantal elementen gerapporteerd worden die noodzaken tot enige terughoudendheid. In wezen komen de kritieken neer op het feit dat aan de hand van zelf- en peerevaluatie moeilijk valide en betrouwbare evaluatiegegevens verzameld kunnen worden. We gaan in de volgende paragraaf eerst na of deze kritiek terecht is. Daarna staan we stil bij een aantal voorwaarden die vervuld moeten zijn om toch tot voldoende kwaliteitsvolle evaluatiegegevens te komen.

2.2.1. De betrouwbaarheid en validiteit bij vormen van leerlingenbetrokkenheid

Het garanderen van de betrouwbaarheid en validiteit van een beoordeling door leerlingen is niet altijd vanzelfsprekend. Wat peerevaluatie betreft, concludeert Topping (1998) uit een vergelijking van 31 studies dat toch voldoende betrouwbare en valide informatie verzameld kunnen worden. Peerevaluatie blijkt in het algemeen wel betrouwbaarder dan zelfevaluatie. Onderzoek leert dat leerlingen op een accurate wijze hun eigen werk kunnen evalueren, maar dat zwakkere leerlingen de neiging hebben om hun prestaties te overschatten (Longhurst & Norton, 1997; Sluijsmans et al., 1999). Een mogelijke verklaring hiervoor zou zijn dat deze leerlingen geen goed beeld hebben van de betekenis van de vooropgestelde criteria en standaarden.

Uit de onzekerheid omtrent de betrouwbaarheid van vormen van leerlingenbetrokkenheid besluiten we dat deze alternatieve evaluatievormen niet zonder meer geschikt zijn voor het uitvoeren van summatieve evaluaties. Momenteel is het inderdaad zo dat zelfevaluatie vaak geen rol speelt bij het nemen van gewichtige beslissingen (zoals selectie, doorstroming, kwalificeren). Net het feit dat mensen de neiging hebben hun eigen talenten milder te beoordelen dan andere evaluatoren (het zogenaamde 'leniency-effect') is hier wellicht mede verantwoordelijk voor. Bovendien zijn zelfbeoordelingen in de praktijk vaak minder specifiek en minder nauwkeurig dan beoordelingen van getrainde evaluatoren (De Rijke & Beemer, 1995). Die redenen komen er eigenlijk op neer dat de betrouwbaarheid van gegevens verkregen aan de hand van een zelf- en peerevaluatie ontoereikend is om gewichtige beslissingen te nemen. Voor de begeleidende en bijsturende functies van evaluatie stelt dit minder problemen.

2.2.2. Sleutelementen voor een succesvolle leerlingenbetrokkenheid

Bovenstaande problemen hoeven niet te impliceren dat men moet afzien van leerlingenbetrokkenheid. Ze moeten wel de aandacht vestigen op het feit dat de doelen die men met de betrokkenheid wenst te bereiken, enkel kunnen worden gerealiseerd indien voldoende getracht wordt de nodige kwaliteitseisen in acht te nemen. Zo moet men de evaluatiemethode integreren in de instructie en dit op een voor de leerlingen positieve, niet bedreigende en attractieve manier. De volgende alinea's zetten een aantal kritische succesfactoren op een rij die de gerapporteerde problemen met betrouwbaarheid kunnen minimaliseren of zelfs vermijden.

Een belangrijke voorwaarde om leerlingen succesvol bij hun evaluatieproces te betrekken is dat ze het gewoon zijn om zichzelf en anderen te evalueren. Cheng & Warren (1997) onderzochten de houding van leerlingen vóór een peerevaluatie en bij het einde ervan. Leerlingen die eerder weigerachtig staan tegenover deze evaluatievorm geven hiervoor volgende redenen aan: ze voelen zich te onervaren om anderen te beoordelen, maar ze zijn ook bang om zelf onfair beoordeeld te worden door hun studiegenoten. Ook de leerlingen twijfelen bovendien aan de ernst en de objectiviteit van de beoordeling. Het is niet vanzelfsprekend dat ze open staan voor zelf- en peerevaluatie. Getuige hiervan reacties van leerlingen zoals 'de leerkracht weet het toch best' en 'ik ben hier om te leren en niet om te onderwijzen' (Zariski, 1996). Om deze weigerachtigheid te overwinnen, moeten twee voorwaarden vervuld worden.

1. Men moet zich in elke evaluatiecontext de vraag stellen in welke mate leerlingen in staat zijn mekaar waardevolle en leerrijke feedback te bezorgen. Belangrijk hiertoe is het creëren van een sfeer van coöperatie in plaats van competitie. Leerlingen zullen zelf enkel eerlijke feedback geven en feedback van medeleerlingen aanvaarden indien ze in het klasgebeuren voldoende vertrouwen, respect en begrip ervaren.
2. Uiterst belangrijk is verder dat er duidelijkheid en eensgezindheid bestaat over de gehanteerde evaluatiecriteria. De leerlingen moeten daarom goed voorbereid worden op hun actieve rol in het evaluatieproces. De kwaliteit van feedback hangt af van het al dan niet succesvolle gebruik van de vooropgestelde criteria. Men moet leerlingen dan ook leren de criteria adequaat te hanteren. Zoniet weten ze niet waarop ze moeten letten bij het evalueren van het werk van medeleerlingen. Hiertoe moeten ze beschikken over de nodige metacognitieve kennis en vaardigheden. Leerlingen moeten bereid zijn om deze nieuwe rol op zich te nemen en om hiervoor de nodige inspanningen te leveren. Het vraagt van hen een mentaliteitswijziging en specifieke vaardigheden. Voorafgaandelijke training is noodzakelijk opdat de leerlingen zich in staat zouden voelen om elkaar op een eerlijke manier te evalueren en om mekaar bruikbare feedback te geven.

3 Het verloop van leerlingenbetrokkenheid

Aanvangen met zelf-, peer- en co-evaluatie is niet vanzelfsprekend. Er moet dan ook voldoende aandacht worden besteed aan de implementatie en het verloop van deze evaluatiemethoden. In dat verloop onderscheiden we een vijftal stappen. De invulling van elk van deze stappen kan verschillen naargelang de aard van de participatie van de leerlingen in het geheel (docentgestuurd, samen, leerlinggestuurd). Hieronder lichten we deze verschillende stappen toe (gebaseerd op Clemens et al, 1999; Dochy & Segers, 1999; Bartels, 1998; Rolheiser, 2000).

Stap 1: Definieren en verkennen van leerlingenbetrokkenheid

De eerste stap bestaat erin te definiëren hoe de leerlingenbetrokkenheid er zal uitzien en waarom deze leerlingactiverende werkwijze belangrijk is voor de leerlingen en hun leerproces. Een te beperkt inzicht in de processen van de betrokkenheid en reflectie kan de bereidwilligheid van leerlingen om aan het evaluatieproces te participeren, belemmeren en kan onnodige obstakels creëren.

Leerlingen moeten in de eerste plaats overtuigd worden om constructief mee te werken aan het evaluatieproces. De weigerachtigheid moet weggenomen worden door leerlingen de kans te geven zich de noodzakelijke vaardigheden eigen te maken en door het creëren van een veilige onderwijscontext. Wanneer leerlingen inzicht hebben in de omschrijving en in het waarom van de leerlingenbetrokkenheid is het belangrijk ze vertrouwd te maken met het eigenlijke evaluatieproces. Alvorens met de zelf- of peerevaluatie aan te vangen, is het daartoe noodzakelijk met de leerlingen een aantal voorbeelden van zelf- en peerevaluaties te bespreken. Hierbij staan

modellen van de vereiste vaardigheden, de rol en het gebruik van criteria, en discussie centraal (Rolheiser et al, 2000). De informatie die leerlingen opdoen door zelfevaluatie van anderen te observeren en te bespreken, zal hen later leiden bij hun eigen zelf- en peerevaluatie. Pas wanneer leerlingen vertrouwd zijn met de verschillende manieren om tot leerlingenbetrokkenheid te komen, is het zinvol eigen criteria te gaan bepalen en te gaan gebruiken.

Ook de doelen die geëvalueerd zullen worden, moeten in de eerste stap geëxpliciteerd worden. Indien mogelijk kan ook hierin een rol worden weggelegd voor de leerlingen. Leerkrachten en leerlingen kunnen bijvoorbeeld samen bediscussiëren welke doelen ze het meest en minst belangrijk vinden. Ook vragen omtrent de wijze waarop geëvalueerd zal worden en het tijdstip van de evaluatie kunnen in overleg bepaald worden. Een bijkomend aspect dat steeds moet bekeken worden in het verloop van een peerevaluatie, betreft het al dan niet anoniem verzamelen van feedback op elkaars werk. De vraag die hier eigenlijk gesteld wordt, betreft dus de invloed van persoonlijkheidskenmerken en interpersoonlijke relaties op de accurateheid van peerevaluaties. Om dergelijke invloeden te vermijden, zou men kunnen opteren voor anonieme feedback. Daartegenover staat dat een groep van leerlingen die elkaar voldoende kent en redelijk hecht is, een goed klimaat vormt voor betrouwbare en valide peerevaluatie (Zariski, 1996).

Stap 2: Het bepalen van criteria

In de tweede fase van het evalueren van leerlingen staat het ontwerpen en vaststellen van evaluatiecriteria centraal. Een degelijke set van criteria kan door de leerkracht en de leerlingen gehanteerd worden om een bepaalde prestatie te evalueren. De manier waarop deze criteria tot stand komen en meer specifiek de betrokkenheid van leerlingen in het ontwikkelen van de criteria kan sterk variëren. Op het continuüm gaande van lage naar hoge leerlingparticipatie kunnen we volgende mogelijkheden onderscheiden: de leerkracht laat de leerlingen in het geheel niet participeren; de leerkracht laat de leerlingen reageren op de criteria die hij opstelde; de leerkracht laat de leerlingen kiezen uit een hoeveelheid criteria; en ten slotte de leerkracht overlegt over de mate van leerlingparticipatie en over het al dan niet hanteren van criteria die door de leerlingen werden vooropgesteld. Essentieel bij leerlingenbetrokkenheid is dat de evaluatiecriteria expliciet kenbaar gemaakt worden en dat de mogelijkheid bestaat om over de criteria in discussie te treden.

Bij collaboratieve evaluatie ligt de nadruk steeds op het gezamenlijk overleg tussen de leerkracht en de leerlingen over de eigenlijke evaluatie. Binnen collaboratieve evaluatie wordt echter een onderscheid gemaakt tussen een zwak en een sterk model. Het onderscheid tussen beide modellen is dus gebaseerd op de verschillen ten aanzien van de manier waarop de evaluatiecriteria tot stand komen:

- ✓ In het **zwakke model** zijn de evaluatiecriteria reeds vastgelegd. Er zijn noch door de leerlingen noch door de leerkracht aanpassingen mogelijk. De leerkracht informeert de leerlingen over de criteria en tracht met hen overeen te komen op welke manier de criteria best geïnterpreteerd worden.

- ✓ In het **sterke model** liggen de criteria nog niet vast. Afhankelijk van het feit of de leerlingen al dan niet in staat zijn zelf criteria te vinden, zijn er twee mogelijke alternatieven. Ofwel stelt de leerkracht een lijst met criteria samen die vervolgens met de leerlingen bediscussieerd en eventueel aangepast wordt. Ofwel kunnen de leerlingen individueel een lijst met criteria aanleggen waarna men doorheen een discussie tot een finale lijst komt.

Wanneer evaluatiecriteria aan de leerlingen bekend gemaakt worden, is het uiteraard belangrijk die af te stemmen op het niveau van de leerlingen. Leerlingen die weten wat van hen verwacht wordt en die de criteria begrijpen alvorens ze een bepaalde opdracht aanvaarden, zullen beter aan die criteria en standaarden beantwoorden. Duidelijkheid over de criteria is dan ook een must. Op termijn is het bovendien nastrevenswaardig dat leerlingen ook eigen criteria ontwikkelen, vooropstellen en nastreven. Op basis van een zelf- of peerevaluatie kunnen leerlingen bijvoorbeeld eigen doelen vooropstellen waarvoor ze zelf aangeven waaraan hun prestaties moeten voldoen om als voldoende kwaliteitsvol te worden bestempeld.

Bij het ontwikkelen en selecteren van criteria is ten slotte volgende opmerking van belang. Bij zelfevaluatie (net als bij peerevaluatie en portfolio's) moet een onderscheid gemaakt worden tussen criteria voor het beoordelen van een bepaalde prestatie (bijvoorbeeld een grafiek maken) en criteria ter beoordeling van de zelfevaluatie zelf. De eerste reeks criteria verschilt in wezen niet van de criteria die leerkrachten gebruiken indien ze leerlingen niet in het evaluatieproces betrekken. De reeks tweede criteria wordt specifiek bepaald om de kwaliteit van de zelfevaluatie te beoordelen. Om dergelijke criteria te illustreren, presenteren we een mogelijke activiteit om zelfevaluatiecriteria te bepalen. Om de kwaliteit van een peerevaluatie te beoordelen kan op een gelijkaardige manier tewerk worden gegaan.

Activiteit 4.1: Bepalen van zelfevaluatiecriteria

Zelfevaluatiecriteria kunnen bepaald worden door de aandacht te richten op reeds bestaande zelfevaluatie. De activiteit bestaat erin individueel of in groepen van een aantal leerkrachten, leerlingen of beide zowel de positieve als negatieve reacties bij concrete zelfevaluatie te beschrijven. Bij wijze van voorbeeld geven we twee korte zelfevaluatievoorbeelden. Bij beide geven we een aantal positieve en negatieve reacties.

Stijn (17): Het valt me op dat ik tegenwoordig vlotter in groep kan werken dan vroeger. Ik voel me meer op mijn gemak. Vroeger had ik de gewoonte bij elk groepswork zeer veel werk te doen. Ik deed dat niet omdat de groep niet wilde meewerken maar omdat ik per se wou dat alles op mijn manier gebeurde. Deze keer heb ik geprobeerd het werk beter te verdelen en ondertussen zelf toch nog veel werk te doen. Ik werkte samen met de rest van de groep in plaats van de groep mijn visie en werkwijze op te leggen. Ik voelde dat mijn klasgenoten het leuker vonden om met mij samen te werken.

- + Reflectie over een tijdsperiode.
- + Stijn had zich een duidelijk doel vooropgesteld.
- + Stijn geeft eerdere fouten toe en durft zich positief over zichzelf uitspreken.

- Het is nog onduidelijk waarin de nieuwe aanpak van Stijn bestaat.
- Wat zijn Stijns toekomstige doelen?

Astrid (16): Deze stageperiode heb ik voor de eerste maal ervaren hoe belangrijk het beroepsgeheim is voor een verzorgende. De bejaarden vragen van alles over mekaar en willen dat ik dat dan vertel. Ik stond vroeger niet stil bij de slechte gevolgen waar-toe die uitspraken zouden kunnen leiden. Daarom probeer ik voortaan zeer voorzichtig te zijn om zaken te vertellen over andere bewoners van het rusthuis.

- + De leerling gaat in op belangrijke leerervaringen.
- + Astrid identificeert strategieën die haar in de toekomst garanderen geen fouten meer te maken. Ze maakt duidelijk dat ze nog verder moet groeien.
- De reflectie is niet echt uitgewerkt (waaruit bestond het leerproces? heeft Astrid voldoende vertrouwen in zichzelf om verder te leren? wat maakt het beroepsgeheim moeilijk om te respecteren? etc.).

Activiteiten als deze kunnen criteria aan het licht brengen die gebruikt kunnen worden om de zelfevaluatie van leerlingen te beoordelen. Nogmaals: dit zijn andere criteria dan de criteria die leerlingen in een zelfevaluatie gebruiken om een bepaalde taak te beoordelen. Volgende lijst zet een aantal mogelijke zelfevaluatiecriteria op een rij. Alle criteria zouden het resultaat kunnen zijn van bovenstaande activiteit.

Figuur 4.33: Zelfevaluatiecriteria

- ✓ Er worden sterke en zwakke punten erkend en verklaard.
- ✓ De leerling heeft een kritische houding ten aanzien van zichzelf.
- ✓ Er worden specifieke strategieën en doelen geïdentificeerd met het oog op toekomstige groei.
- ✓ De leerling gaat in op relevante leerprocessen (er is aandacht voor product en proces).
- ✓ De zelfevaluatie wordt voldoende uitgewerkt, is gedetailleerd.
- ✓ De leerling geeft een verklaring voor de inhoud van de reflectie.
- ✓ Het is duidelijk dat de betrokkene zelf aan het woord is (ervaringen, gevoelens, plannen, etc.).
- ✓ De reflectie is eerlijk en oprecht.
- ✓ De reflectie is goed gestructureerd en leesbaar.

Stap 3: Het toepassen van de criteria – de eigenlijke beoordeling

Wanneer men over een goede batterij criteria beschikt die de basis kunnen vormen van een evaluatie, komt het erop aan leerlingen deze criteria te leren toepassen. Dit leerproces is gelijkwaardig aan gelijk welk ander leerproces dat tot doel heeft leerlingen bepaalde vaardigheden bij te brengen. Afhankelijk van de mate van participatie van de leerlingen en de feedback die men leerlingen geeft, kan deze fase er echter zeer verschillend uitzien. We pleiten ervoor leerlingen zoveel mogelijk feedback te geven wanneer ze leren de evaluatiecriteria toe te passen. Op basis van concrete voorbeelden en levensechte toepassingen kunnen leerlingen zich het gebruik van criteria eigen maken.

Enmaal leerlingen over de noodzakelijke vaardigheden beschikken om evaluatiecriteria te gebruiken, wordt het mogelijk om effectief beoordelingen uit te spreken. De criteria en beoordelingsschalen kunnen met andere woorden van dan af gebruikt worden om (eigen) prestaties te beoordelen. Bij het definiëren van de verschillende vormen van alternatieve evaluatie die de betrokkenheid van leerlingen nastreven, stelden we dat net in de beoordelingsfase het onderscheid tussen zelf-, peer- en co-evaluatie duidelijk wordt. Dat blijkt uit de verschillende manieren waarop leerlingen kunnen worden betrokken in het gebruik van evaluatieinstrumenten. We richten ons in figuur 4.34 ter illustratie op het beoordelen van een door de leerlingen gemaakte grafiek.

Figuur 4.34: Een instrument voor zelf-, peer- en co-evaluatie

	1	2	3	4
De titel van de grafiek geeft duidelijk het verband aan tussen de assen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Er werd een geschikt grafiektype gebruikt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
De lijnen of balken gebruiken de beschikbare ruimte goed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beoordelingsschaal: 1=onvoldoende; 2= voldoende met leemtes; 3= goed; 4=zeer goed				

Een dergelijk instrument kan door verschillende beoordelaars gebruikt worden. Bij zelfevaluatie gebruiken leerlingen het om een zelfgemaakte grafiek te beoordelen. Bij peerevaluatie hanteren leerlingen hetzelfde instrument om grafieken te beoordelen die door hun klasgenoten gemaakt werden. Of leerlingen kunnen een eigen grafiek tezamen met de klasgenoten evalueren. Hiertoe kunnen verschillende technieken gebruikt worden. Wanneer descriptieve schalen gebruikt worden, kan bijvoorbeeld gezocht worden naar een unanieme beoordeling op de verschillende criteria. Twee of meer leerlingen bespreken de kenmerken van de grafiek en beargumenteren waarom die in hun ogen een bepaalde beoordeling verdient. Op die manier kan de peerevaluatie uiteraard niet anoniem gebeuren. Dat kan wel door de zelfevaluatie van een leerling te confronteren met de anonieme beoordeling van een klasgenoot. Ook op die manier worden leerlingen uitgenodigd om te reflecteren over de kwaliteit van de eigen prestatie.

Wanneer de beoordeling van een product of een prestatie gebaseerd is op een waardering van meerdere betrokkenen doet men vaak een beroep op numerieke scoringsmethoden. Met getallen kan op een creatieve en interessante manier omgegaan worden. Stel dat je de bijdrage van een individuele leerling aan een groepswork wil evalueren, dan kan je bijvoorbeeld op volgende manieren te werk gaan (Dochy & Segers, 1999).

a. Vermenigvuldigen van een groepscoëfficiënt met een individuele wegingsfactor

Het groepscoëfficiënt is gebaseerd op de beoordeling door de leerkracht voor het werk van de groep in haar geheel. Als ik een beoordeling op een 100-puntenschaal wil uitspreken, is dit cijfer bijvoorbeeld gelijk aan 75. Het individuele beoordelingscijfer is dan gelijk aan het groepscoëfficiënt van de leerkracht vermenigvuldigd met een beoordelingscijfer van de

medeleerlingen. Dat is de peerevaluatie-factor. Het wordt bekomen door elke individuele leerling op een aantal criteria te beoordelen aan de hand van de volgende schaal:

- (3) beter dan de rest van de groep
- (2) gemiddeld (in verhouding tot de groep)
- (1) niet zo goed als de rest van de groep
- (0) geen hulp voor de groep
- (-1) hinder voor de groep

Elke leerling van de groep spreekt een oordeel uit over de eigen prestaties en de prestaties van de andere groepsleden. Dat leidt bijvoorbeeld tot volgend resultaat.

	Jezelf	Pieter	Alex	Ingrid
Vorbereiding van individueel werk	2	-1	3	2
Aanreiken van ideeën	2	1	3	2
Inzicht in de materie	2	2	2	2
Bewaken van de positieve relaties	1	2	1	2

Voor elke leerling wordt het gemiddelde berekend van de verschillende beoordelingen. Deze worden afgewogen tegen het gemiddelde van de groep (dat is 2). De individuele peerevaluatie-factor wordt berekend door de eigen gemiddelde score te delen door de gemiddelde groepsscore. Bij Alex is deze score gelijk aan 1.125 ($= (3+3+2+1)/4$ gedeeld door 2). Voor Ingrid is deze score gelijk aan 1 (ze scoort telkens gemiddeld). Wanneer een peerevaluatie-factor gebaseerd is op het oordeel van verschillende groepsleden, dan neemt men het gemiddelde van de verschillende peerevaluatie-factoren. Door deze peerevaluatie-factor vervolgens te vermenigvuldigen met het groeps cijfer bekom je de co-evaluatiescore. Voor Alex is deze gelijk aan 84.5 ($= 75 \times 1.125$) en voor Ingrid aan 75 ($= 75 \times 1$). Op een gelijkaardige manier kan je een zelfevaluatiefactor berekenen en vermenigvuldigen met het groeps cijfer.

Een mogelijke variant bestaat erin het groeps cijfer te verhogen of te verlagen met de beoordeling van de peers. Een leerling die volgens zijn medeleerlingen gemiddeld meewerkte, behoudt zo het groepsgemiddelde terwijl andere leerlingen als het ware bijkomende punten krijgen van hun medeleerlingen dan wel punten worden afgenomen.

b. Verdeling van een cijfer

Stel dat leerlingen in groepjes van vier samengewerkt hebben aan het schrijven van een brochure of tezamen een fietsrek hebben ontworpen. De leerkracht geeft bij deze werkwijze de groep een bepaald beoordelingscijfer, bijvoorbeeld 70%. Hieruit volgt dat in een groep van vier elk lid in zijn peerevaluatie 280 punten kan verdelen ($= 4 \times 70$ punten). Na het gezamenlijk bespreken van de vooropgestelde criteria en eventueel na het gebruik van rubrieken en evaluatielijsten, spreken leerlingen over mekaar beoordelingen uit. Dat kan bijvoorbeeld op volgende manier.

	Bram	Hendrik	Stefaan	Celcuck
Bram geeft aan	80	60	65	75
Hendrik geeft aan	75	60	65	80
Stefaan geeft aan	85	55	60	80
Celcuck geeft aan	75	65	65	75
Gemiddelde score	79	60	64	77.5

Leerlingen kunnen op verscheidene manieren betrokken worden in het eigenlijke beoordelingsproces. De eigenlijke methodes van zelf-, peer- en co-evaluatie kunnen dus verschillende vormen aannemen. Naargelang de kenmerken van de leerlingen en de doelen van de evaluatie moet steeds naar een geschikte werkwijze gezocht worden.

Stap 4: Het geven van feedback

Het voornaamste doel van leerlingenbetrokkenheid is het stimuleren van leerlingen tot reflectie. Net in de terugkoppelingsfase kunnen leerlingen hun eigen sterke en zwakke punten aan het licht brengen en plannen maken voor verbetering. Het geven van feedback is daarom een proces dat zoveel mogelijk in een onderlinge samenwerking tussen de leerling zelf, de medeleerlingen, de leerkracht en eventuele derden moet plaatsvinden.

De mate van betrokkenheid van leerlingen kan ook bij het geven van feedback echter sterk variëren. Dit ongeacht de mate waarin leerlingen al dan niet betrokken werden in de vorige fasen van het evaluatieproces. Als leerkracht kan je ervoor opteren de leerling enkel de feedback te geven die je zelf wenst te geven. De leerling is dan een passieve ontvanger van feedback. Meer betrokkenheid is er wanneer de leerling de mogelijkheid krijgt de eigen zelf- of peerevaluatie te verantwoorden of wanneer de leerkracht de leerling in een dialoog betreft over de feedback op de evaluatie. Het verloop van het geven van feedback is dus sterk afhankelijk van wie er feedback geeft. Ofwel wordt die enkel door de leerkracht verzorgd ofwel kunnen medeleerlingen ook feedback geven. Uiteraard kunnen ook andere bronnen van feedback aangewend worden (zoals buitenstaanders, stagementoren, etc.).

Stap 5: Het stellen van leerdoelen

Eenmaal er gereflecteerd werd op het leren, kan men nieuwe doelstellingen vooropstellen. Wanneer de nieuwe leerdoelen gebaseerd worden op de resultaten van een zelf-, peer- of co-evaluatie, kan het verloop van deze stap variëren al naargelang de mate waarin leerlingen hierbij betrokken worden. In de afwezigheid van leerlingparticipatie is enkel de leerkracht verantwoordelijk voor het bepalen van de leerdoelen. Het gevaar kan dan bestaan dat de leerdoelen niet altijd even goed zijn afgestemd op de specifieke kwaliteiten van de betreffende leerling. Een zekere garantie hiertoe biedt het betrekken van de leerlingen bij het stellen van de doelen. De leerkracht kan bijvoorbeeld op basis van de uitkomsten van de zelfevaluatie een aantal mogelijke leerdoelen vooropstellen waaruit samen met de leerling die doelen weerhouden worden die het best afgestemd zijn op de eigen kwaliteiten van de leerling. De leerlingparticipatie kan nog versterkt worden door de leerling -al dan niet op basis van de zelfevaluatiegegevens- eigen doelen te laten formuleren. Deze doelen kunnen zowel betrekking hebben op het leerproces als op het leerproduct.

Men zou het stellen van nieuwe leerdoelen kunnen aanzien als de eerste stap van een nieuwe evaluatie. Na het stellen van leerdoelen moeten immers opnieuw criteria bepaald worden. De verschillende stappen hebben dan ook een cyclisch karakter.

4 Evaluatieinstrumenten

De evaluatieinstrumenten die aangewend kunnen worden wanneer men leerlingen actief wil betrekken bij hun eigen evaluatie, zijn verscheiden. Afhankelijk van de doelen van de evaluatie enerzijds en de hoeveelheid en de aard van de sturing die leerlingen krijgen anderzijds zal de zelfevaluatie er telkens anders uitzien. De volgende paragrafen zetten een aantal werkwijzen op een rij die tot doel hebben leerlingen te laten reflecteren over hun eigen leren. Hieruit zal blijken dat het er in een aantal gevallen op neerkomt dat niet de leerkracht maar de leerlingen zelf beoordelingsinstrumenten gaan gebruiken. In andere gevallen worden instrumenten speciaal ontwikkeld om leerlingen te stimuleren tot reflectie.

4.1 Evaluatielijsten

Een veel gebruikte werkwijze om leerlingen te laten reflecteren, bestaat uit het aanbieden van vragen(lijsten) die leerlingen stimuleren om over hun eigen leerproces en producten na te denken. Bij het beschrijven van het evaluatieproces (zie deel 3) werd beschreven hoe beoordelingsinstrumenten vorm kunnen krijgen. Vooral het gebruik van evaluatielijsten werd benadrukt. Een evaluatielijst heeft als doel de beoordelaars van een leerlingprestatie bij te staan in het beantwoorden van de vraag wat de kwaliteitsaspecten in een bepaald product of in een bepaalde leerlingprestatie zijn (ASCD, 1996). Evaluatielijsten kunnen aldus eenvoudig worden aangewend als instrumenten om de leerlingprestaties door de leerlingen zelf te laten beoordelen. Net als leerkrachten kunnen leerlingen immers aan de hand van dergelijke instrumenten beoordelen of, en in welke mate, de eigen prestatie voldoet aan de criteria die voorop werden gesteld. De-

zelfde instrumenten kunnen worden aangewend om de prestaties van medeleerlingen te beoordelen.

Figuur 4.35: Zelfevaluatielijsten geïllustreerd

De volgende zelfevaluatielijst heeft tot doel leerlingen Personenzorg te laten reflecteren over de manier waarop ze er in slagen samen met kinderen te spelen. Bij wijze van voorbeeld gebruiken we de volgende kwalitatieve evaluatieschaal: mijn gedrag is een voorbeeld voor anderen; dit kan ik; hier moet ik nog wat aan werken; hier schiet ik werkelijk tekort. Leerlingen moeten zichzelf op deze schaal situeren bij elk van de volgende criteria:

- ✓ Vragen stellen om de belangstelling van het kind te achterhalen.
- ✓ Verwondering weten op te wekken.
- ✓ Kunnen bekennen iets niet te weten of niet te kunnen.
- ✓ Meedoen aan gemeenschappelijke activiteiten.
- ✓ Verantwoordelijkheid geven aan kinderen.
- ✓ Initiatieven van de kinderen stimuleren.
- ✓ Met kinderen samen mogelijkheden bekijken, keuzes afwegen en beslissingen nemen.
- ✓ Je gedrag als verzorgende bespreekbaar durven maken.

Kenmerkend voor het gebruik van dergelijke zelfevaluatielijsten is het feit dat de evaluatie nauw aansluit bij de criteria die werden vooropgesteld. Leerlingen worden bovendien sterk gestuurd. Dat is evenwel niet steeds noodzakelijk. Naast het gebruik van evaluatielijsten kunnen dan ook nog andere vormen van zelfevaluatievragenlijsten worden onderscheiden, elk met hun eigen doelen en doelgroep. De volgende selectie voorbeelden van vragenlijsten is indicatief voor de variatie en creativiteit die zelfevaluatievragenlijsten kunnen kenmerken.

- ✓ Het aanbieden van zinnen die leerlingen moeten aanvullen, kan de reflectie op de eigen prestaties stimuleren en sturen. Op die manier kan een verscheidenheid aan voor het leerproces relevante elementen aan bod worden gebracht.

Aanvullen van zinnen

- ✓ Ik ben zeer tevreden over...
- ✓ Wat ik voel bij wat ik bereikt heb, is...
- ✓ Ik moet nog werken aan...
- ✓ Wat ik geleerd heb van de vorige opdracht(en) is...

Dergelijke aanvulzinnen kunnen nog opener geformuleerd worden:

- ✓ Wat was voor jou het moeilijkste aspect van deze opdracht?
- ✓ Wat ga je volgende keer anders doen?
- ✓ Hoe denk je dat best aan te pakken?

- ✓ Leerlingen die nog onervaren zijn met zelfevaluatie kunnen zich aanvankelijk aarzelend opstellen. Het komt er dan ook op aan met kleine zelfevaluatie- en reflectiemomenten aan te vangen en deze systematisch uit te breiden. Probeer hierbij interessante en motiverende zelf- en peerevaluatievragen aan te reiken (Mabry, 1999).

Korte opdrachten

Beschrijf in één woord (een adjectief) of in één zin in welke mate je tevreden bent over je eigen prestaties.

Creatieve opdrachten

Kies een titel van een liedje, film, boek of een stripfiguur die het best weergeeft hoe je denkt over je eigen werk. Leerlingen kunnen hierbij in een aantal zinnen hun keuze verantwoorden. Denk bijvoorbeeld aan de volgende voorbeelden:

- ✓ I get by with a little help from my friends
- ✓ Help
- ✓ The talented Mr. Ripley
- ✓ Nobelprijs
- ✓ Piet Pienter

Beoordelingsschalen kunnen op verschillende manieren gevisualiseerd worden. Zo kan men leerlingen het aantal graden laten aanduiden op een thermometer of op een dartsbord laten aangeven hoe ver ze zich nog van de roos situeren.

- ✓ Jonge leerlingen of leerlingen die problemen hebben met lezen en schrijven kunnen gevraagd worden mondeling of visueel hun beoordeling van de kwaliteit van hun eigen werk weer te geven.

Visuele zelfevaluatie

Teken de rest van dit gezicht zodat het aantoont hoe jij je voelt bij jouw werk. Of als variatie: selecteer het gelaat dat best jouw beoordeling van je werk weergeeft



4.2 Attitudeschalen

Een specifieke vorm van zelfevaluatielijsten zijn attitudeschalen. Ook dat zijn zelfbeschrijvende technieken. Ze richten zich op houdingen en opinies van de leerlingen. Het verschil met andere evaluatielijsten is dat attitudeschalen enkel door de te evalueren leerling zelf kunnen worden gebruikt. Ik kan dus zelf bijvoorbeeld geen attitudeschaal invullen om uitspraken te kunnen doen over de houding van mijn klasgenoot ten opzichte van het vak natuurkunde.

Om houdingen of opinies te evalueren, wordt meestal een Likert-schaal gebruikt. Dat is een schaal die bestaat uit meerdere verwante items (uitspraken of stellingen) waarop leerlingen verzocht worden te reageren door het kiezen voor één der voorgegeven antwoordmogelijkheden (Thio, 1988):

- | | |
|---------------------|---------------------------|
| (5) zeer mee eens | (5) volledig akkoord |
| (4) mee eens | (4) akkoord |
| (3) neutraal | (3) neutraal |
| (2) mee oneens | (2) niet akkoord |
| (1) zeer mee oneens | (1) helemaal niet akkoord |

Het gebruik van attitudeschalen illustreren we aan de hand van de houding van leerlingen ten opzichte van de medeleerlingen (zie figuur 4.36 - gebaseerd op Van Putten, 1987).

Figuur 4.36: Attitudeschaal 'relaties met medeleerlingen'

	Dat is zo	Soms wel, soms niet	Dat is niet zo
Ik heb in mijn klas weinig vrienden of vriendinnen	O	O	O
Ik vind het moeilijk om in mijn klas vrienden of vriendinnen te maken	O	O	O
Ik vind het leuk om met mijn klasgenoten om te gaan	O	O	O
Ik voel mij in onze klas er een beetje buiten staan	O	O	O
Ik voel mij goed op mijn gemak bij mijn klasgenoten	O	O	O
Ik vind mijn klasgenoten niet zo aardig	O	O	O
Ik kan met de meeste van mijn klasgenoten goed opschieten	O	O	O

Het ontwerpen van dergelijke schalen is niet eenvoudig. Dat neemt niet weg dat attitudeschalen zeer interessante informatie kunnen opleveren over de houding van de leerling (Standaert & Troch, 1998). Attitudes van leerlingen kunnen immers een belangrijke invloed uitoefenen op de prestaties van leerlingen. Denken we maar aan elementen als motivatie, faalangst, examenangst, concentratievermogen, etc. De gegevens die aan de hand van dergelijke vragenlijsten bekomen worden, zijn dus zeer interessant om de prestaties van leerlingen te begrijpen. Daartegenover staat dat de betrouwbaarheid van deze gegevens vaak bekritiseerbaar is. Twee redenen zijn hiervoor verantwoordelijk. Enerzijds is het zeer moeilijk betrouwbare attitudeschalen op te stellen. Het is dan ook aangewezen om die elders te ontleen. Anderzijds moet de nodige voorzichtigheid aan de dag worden gelegd omdat leerlingen wel eens positievere antwoorden durven geven dat de werkelijkheid zou toelaten (Standaert & Troch, 1998). Het is dan ook niet de bedoeling om louter op basis van een attitudeschaal te beslissen of leerlingen de correcte houdingen hebben om bijvoorbeeld de overgang te kunnen maken naar een volgend schooljaar of om een bepaald certificaat te verdienen. Gegevens uit attitudeschalen moeten daarom worden aangevuld met gegevens uit andere evaluatielijsten, logboeken en gesprekken met de leerlingen.

4.3 Logboeken

Een logboek bestaat uit een continue weergave van indrukken, ervaringen en reflecties bij een bepaald onderwerp (MMS, 2001). Het is als het ware een weergave van de innerlijke dialoog van de leerling. De vergelijking met een dagboek is dan ook niet veraf (ook al hoeft een logboek niet dagelijks te worden bijgewerkt). Het voornaamste doel van een logboek is de leerlingen te stimuleren om actief na te denken over hun eigen leerproces en om hen also hun eigen leerproces in handen te laten nemen. Ook het louter bijhouden en plannen van wat de leerlingen doen, is vaak een belangrijk doel. Het doel is dus niet zozeer een zware kritische analyse of een heus onderzoeksrapport samen te stellen. Een logboek wil leerlingen de kans geven gedachten, impressies en gevoelens naar boven te brengen, die anders geen aandacht dreigen te krijgen. De feedback die leerlingen later bij hun logboek krijgen, maakt de aandacht voor deze elementen nog explicieter.

Om duidelijk te maken wat een logboek is, presenteren we een beschrijving die gebruikt kan worden om de leerlingen er zelf mee te laten kennismaken.

Figuur 4.37: Een toelichting voor leerlingen

Wat is een logboek?

In een logboek beschrijf je regelmatig (bijvoorbeeld als huiswerk na elke les of aan het einde van een dagdeel) wat je gedaan hebt en hou je verscheidene zaken bij: wanneer en waar je iets gedaan hebt; hoe je hebt gewerkt (en je rol daarin); met wie je hebt gewerkt; wat je ervan vond (problemen, oplossingen, wat ging goed, wat heb je geleerd, wat wil je nu leren, wat doe je de volgende keer anders of beter?). Een logboek is persoonlijk en daarom leg je het doorgaans alleen aan. Iedereen zal andere voorwerpen de moeite waard vinden om te schetsen, en heeft een andere schrijfstijl. Het kan ook zeker geen kwaad wanneer jij een andere visie hebt dan je leerkrachten.

Waarom moet je een logboek maken?

Een logboek heeft twee functies: het is voor jou een hulpmiddel bij het systematisch verzamelen van informatie en voor je leerkracht is het een hulpmiddel om snel inzicht te krijgen in het gevolgde werkproces. Door een logboek te maken leer je gestructureerd te werken. Doordat je steeds noteert hoe en waarom je tot bepaalde keuzes komt, dwing je jezelf over die keuzes na te denken. Daardoor wordt je werkproces en daarmee het eindproduct beter. Ook leer je op deze manier beter zelfstandig tot keuzes te komen. Daarnaast is het ook voor de planning heel belangrijk: in een logboek kan je heel gericht eigen doelstellingen formuleren, je weet dan precies wat je binnen een bepaalde tijdsperiode moet doen.

Een standaardlogboek bestaat uit een (chronologische) weergave van gebeurtenissen en opinies die het de leerling mogelijk maken te reflecteren over een verscheidenheid van binnen- en buitenschoolse leerervaringen (lezen, schrijven, observeren). De manieren waarop leerlingen die ervaringen en reflecties weergeven, kunnen in principe zeer verscheiden zijn. Een logboek kan zowel worden neergeschreven als worden ingesproken op een bandje. Het bijhouden van een logboek hoeft niet dagelijks te gebeuren maar het lijkt wel aangewezen. Men kan als leerkracht hiervoor bijvoorbeeld dagelijks 10 of 15 minuten tijd reserveren.

Omdat deze vorm van zelfevaluatie de leerling uitnodigt op een zeer persoonlijke en intieme manier ervaringen te expliciteren, vereist het gebruik van logboeken empathische en authentieke leerkrachten en een voor de leerlingen veilige omgeving. Dit betekent bijvoorbeeld dat logboeken best niet in de vorm van cijfers of graden beoordeeld worden. De inspanningen van de leerling en de betekenis die de inhoud van het logboek voor de leerling heeft, staan centraal. Leerlingen mogen niet het gevoel hebben te moeten presteren of de gevolgen van een negatieve beoordeling te moeten vrezen. Leerlinggerichte beoordelingstechnieken lijken dus aangewezen. De persoonlijke inbreng van de leerlingen impliceert bovendien dat een doordachte organisatie en structuur noodzakelijk zijn zodat het logboek een legitiem en integraal onderdeel kan uitmaken van het leerproces. Daarnaast moeten de leerkracht en de leerlingen duidelijke afspraken maken over de manier waarop met zeer persoonlijke opmerkingen wordt omgegaan.

4.3.1. Verschillende types logboeken

Naargelang de sturing die aan leerlingen gegeven wordt en de inhoud ervan, kunnen verschillende types logboeken onderscheiden worden. De volgende paragrafen verduidelijken deze verschillen.

1. Verschillen naar sturing

Naargelang het doel van het logboek en de kenmerken van de leerlingen die er gebruik van zullen maken, wordt een onderscheid gemaakt tussen weinig gestructureerde en voorgestructureerde logboeken.

a. Een weinig gestructureerd logboek

Leerlingen schrijven, met een minimale sturing, wat ze denken en voelen. De leerlingen moeten zelf nog talrijke beslissingen nemen over de inhoud en de structuur van hun logboek. Bij ervaren leerlingen kan de opdracht bijvoorbeeld zonder meer als volgt luiden: houd een logboek bij. Een iets meer gestructureerde opdracht zien er dan weer als volgt uit.

Figuur 4.38: Een stagelogboek

In een logboek beschrijf je elke dag wat je gedaan hebt tijdens de stage, wat er goed ging en wat er minder goed ging (en waarom). Je geeft ook aan welk doel je voor de volgende dag vooropstelt. Het gaat hierbij niet alleen om feiten, maar juist om jouw ideeën, jouw visie op je eigen werk en je leerproces. Wanneer je alleen maar feitelijk beschrijft wat je de stagedag gedaan hebt, is dat onvoldoende. Geef ook heel duidelijk aan waarom je op een bepaalde manier denkt over wat je hebt gedaan en wat je wil gaan doen.

b. Een voorgestructureerd logboek

Leerlingen worden verwacht een specifieke opdracht uit te voeren. Het logboek richt zich op een aantal thema's die door de leerkracht geselecteerd werden (eventueel in samenspraak met de leerlingen). De leerlingen hebben hierbij de nodige afbakeningen en richtlijnen ter beschikking die het mogelijk maken de opdracht succesvol uit te voeren. In figuur 4.39 ge-

ven we een voorbeeld van een voorgestructureerd logboek. Het kan gebruikt worden bij een verscheidenheid aan leer- en evaluatietaken. De verschillende vragen worden in een vastliggend schema aan de leerlingen gepresenteerd.

Figuur 4.39: Een voorgestructureerd logboek

Dit logboek werd ontworpen om je te helpen reflecteren over je huidige leerproces en activiteiten bij deze taak. Hierbij richten we ons meer bepaald op je eigen denkprocessen en je eigen leervaardigheden. Het logboek is chronologisch onderverdeeld in drie delen: wat er gebeurt voor je de taak aanvat en wat er gebeurt tijdens en na het oplossen van de taak. Je moet het logboek dan ook driemaal ter hand nemen en telkens de gestelde vragen beantwoorden.

Voor het oplossen van de taak:

- ✓ Versta je wat er van jou verwacht wordt?
- ✓ Heb je alle informatie om de taak op te lossen?
- ✓ Hoe ga ik deze taak aanpakken?

Tijdens het oplossen van de taak:

- ✓ Wat heb je tot nu toe gedaan?
- ✓ Verloopt alles naar wens? Klopt je planning nog?
- ✓ Is dit de meest geschikte aanpak?
- ✓ Kan je hulp gebruiken?
- ✓ Heb je een goed gevoel bij je huidige werkwijze? Waarom (niet)?

Na het oplossen van de taak:

- ✓ Welke activiteiten heb je ondernomen?
- ✓ Wat ging goed?
- ✓ Wat vond je moeilijk?
- ✓ Wat ga je de volgende keer anders doen?
- ✓ Wat heb je met deze taak geleerd?

2. Verschillen naar inhoud

Logboeken variëren niet alleen naargelang de sturing die leerlingen wordt aangeboden. Ook de inhoud ervan kan sterk verschillen. In principe is het mogelijk een inhoudelijk onderscheid te maken tussen het leren van leerlingen, de taken van leerlingen en de ervaringen en gevoelens van leerlingen. In de praktijk komen echter veelal combinaties van deze inhoud voor.

a. Het leren van leerlingen – learner report

Logboeken hebben vaak tot doel de eigen leerervaringen van leerlingen onder woorden te brengen. Doorgaans zijn de leerervaringen die men verwacht voor alle leerlingen gelijkwaardig. Toch zijn er taken en situaties die kunnen leiden tot onverwachte en/of zeer individueel gekleurde leerervaringen. In die gevallen richt men zich in een logboek specifiek op het leren van leerlingen. Men spreekt dan van een 'learner-report'. Doel hiervan is alle vormen van leren aan de orde te brengen (Clemens et al, 1999). Dat kan aan de hand van volgende soorten vragen:

1. Volledig open vragen: 'Ik heb geleerd dat...' en 'Ik heb me gerealiseerd dat...'.

2. Gespecificeerde vragen: 'Over het verband tussen de aanwezigheid van water en het ontstaan van steden in de Middeleeuwen heb ik geleerd dat...'.
3. Voorgestructureerde vragen: de leerlingen krijgen een lijst met mogelijke leerervaringen aangeboden en moeten op een beoordelingsschaal aangeven in welke mate de bewuste leerervaringen op hen van toepassing zijn.

b. De taken van leerlingen

Logboeken kunnen zich ook toeleggen op de dagelijkse activiteiten van leerlingen. In dat geval vormen ze eerder een soort agenda. De leerlingen geven niet alleen aan welke activiteiten ze tijdens een bepaalde periode uitgevoerd hebben, ze blikken ook vooruit en plannen toekomstige activiteiten. Leerlingen denken na over welke activiteiten ze moeten ondernemen om een bepaalde taak tot een goed einde te brengen. Hierbij geven ze duidelijk aan welke activiteit ze wanneer zullen volbrengen.

c. De ervaringen en gevoelens van leerlingen

Logboeken kunnen zich ook richten op ideeën, vragen, inzichten, dromen en gevoelens van leerlingen. In feite komt het erop neer dat de leerling zelf bepaalt over welke onderwerpen hij schrijft. Alles wat de leerling de moeite waard vindt, is voldoende relevant om te worden opgenomen. Het kan hierbij zowel gaan om zeer persoonlijke als algemene zaken. Een dergelijk logboek sluit nauw aan bij wat we doorgaans kennen als een dagboek.

We stelden dat in de praktijk logboeken veelal een combinatie zijn van deze drie inhoud. Dat blijkt uit het voorbeeld in figuur 4.40.

Figuur 4.40: Een voorbeeld van een logboekbijdrage

De volgende beschrijving is overgenomen uit het logboek van Sandra, een leerling uit de derde graad BSO -Verzorging. Ze zet voor elk van de gezinnen waar ze bij de dienst 'Familiehulp' stage liep de belangrijkste leerervaringen op een rij. Hier rapporteert ze over haar hulp aan Corry, een 78-jarige vrouw.

"Mij is hier vooral opgevallen dat ze nog vrij zelfstandig is. Ze woont in een groot huis met alle comfort. Corry heeft een zeer nauwe band met haar zoon, deze is gehandicapt.

Ik heb hier leren ramen poetsen met in het water azijn om heel propere ramen te krijgen. Ook heb ik leren koken met diepvriesgroenten en maïzena express in plaats van met een roux saus te maken. Ik heb ook vanillepudding leren maken met een ei eronder geklopt zodat dit luchtiger wordt. Ik heb hier leren haar inrollen wat ik vrij interessant vond, omdat dit steeds terugkomt dat ze ernaar vragen. Ik weet wel van mezelf dat ik nog moet leren van het eten allemaal gelijk klaar te hebben.

Het contact met mijn collega was heel goed, ik kon alles aan haar vragen. Ik kreeg op alle vragen een antwoord. Als er iets was, zei ze dit dadelijk. Het contact met Corry was ook zeer goed. Soms had ik moeilijkheden om haar te verstaan, andere moeilijkheden waren er niet bepaald. Soms had ik het wel moeilijk om op bepaalde gesprekken in te gaan over haar zoon. Corry vroeg me veel over mijn vorige stageplaats. Veel kon ik niet vertellen omdat ook daar het beroepsgeheim geldt."

4.3.2. Logboeken: stappen tot reflectie

Net als bij andere zelfevaluatieinstrumenten staat bij logboeken reflectie centraal. Bij logboeken zijn de leerlingen echter nog meer dan elders zelf aan het woord. Om die reden gaan we even dieper in op een kader dat leerlingen kan ondersteunen in hun zelfevaluatie en reflectie (Rolheiser et al., 2000). Dat kader benadrukt twee belangrijke stappen in een zelfevaluatieproces die noodzakelijk zijn om tot een goede reflectie te komen, namelijk vertellen en relateren. De volgende figuur zet een aantal activiteiten binnen verschillende reflectiefases op een rij. In de rechtse kolom worden telkens een aantal aanzetten gegeven om leerlingen op weg te helpen.

Figuur 4.41: Navertellen, relateren en reflecteren

Navertellen	Een aantal aanzetten
✓ In je eigen woorden vertellen	✓ Ik merkte op dat...
✓ Essentiële delen beschrijven	✓ Ik genoot vooral van...
✓ Samenvatten	✓ Ik werd vooral uitgedaagd door...
✓ Ideeën opsommen	✓ De essentie was...
✓ Verduidelijken en details benadrukken	✓ Het meest interessante aspect is...
✓ Specifieke vragen beantwoorden	
Relateren	Een aantal aanzetten
✓ Verbanden leggen en refereren (ervaringen, boeken, gebeurtenissen, ...)	✓ Dat doet me denken aan...
✓ Vergelijken en contrasteren	✓ Dat herinnert me aan...
✓ Analyseren	✓ Dat geeft me een gevoel van...
✓ Illustreren	✓ Dat blijkt uit...
	✓ Ik bedoel hiermee dat...
Reflecteren	Een aantal aanzetten
✓ Conclusies trekken	✓ Nu begrijp ik dat...
✓ Beoordelen en opinies weergeven	✓ Ik vraag me af of ik...
✓ Nieuwe inzichten toelichten	✓ Nu besef ik dat...
✓ Vragen stellen	✓ Ik leerde dat...
✓ Redenen geven voor wat je deed	✓ Misschien...
	✓ Toch denk ik dat...

Om tot een succesvolle reflectie te komen, moeten leerlingen eerst kunnen navertellen en relateren. Dat geldt niet alleen voor geschreven reflecties. Ook in mondelinge interviews moet deze volgorde gerespecteerd worden.

4.4 Mondelinge interviews

Een sterke manier om te reflecteren over de eigen prestaties is erover praten. Leerlingen vinden het over het algemeen plezierig om te praten over hun eigen manier van werken en over hoe ze in de klas functioneren (Clemens et al, 1999). Tenminste als ze in een veilige context op een open manier kunnen praten. Leerlingen moeten elkaar vertrouwen en belangstelling hebben voor elkaar. Bovendien moeten ze beschikken over een aantal communicatieve vaardigheden.

Interviews houden bijvoorbeeld in dat je als leerkracht de leerlingen bevraagt of dat de leerlingen elkaar in tweetallen interviewen. Men spreekt in dergelijke gevallen van interview-interviews.

Om interviews tussen leerlingen te structureren, kan je als leerkracht concrete leidraden opstellen. Dat resulteert bijvoorbeeld in de volgende uitwerking.

Figuur 4.42: Leidraad voor een peerinterview

De leerlingen vormen tweetallen en interviewen elkaar. De inhoud van de interviews wordt eerst door de leerkracht ingeleid. De evaluatiecriteria worden toegelicht en de leerkracht legt de bedoeling van de evaluatiemethode duidelijk uit. De leerlingen worden overtuigd van de zinvolheid van de activiteit. Daarna beginnen de interviews (ongeveer 15 min.). Leerlingen krijgen hierbij een aantal richtvragen, bijvoorbeeld:

- ✓ Wat wilde je met deze prestatie (of dit product) aantonen?
- ✓ Is je dat gelukt?
- ✓ Waaruit leid je dat af?
- ✓ Hoe kan je dat verklaren?
- ✓ Wat leer je hieruit voor later?

Deze vragen geven een duidelijke structuur maar bieden ook voldoende ruimte tot eigen inbreng. De leerlingen maken van elkaars antwoorden korte notities die later uitgewisseld worden. Hierbij wordt sterk de nadruk gelegd op toekomstige doelen en voornemens van de leerlingen. Eventueel kan de leerkracht afsluitend een aantal vragen klassikaal overlopen. Na een gegeven tijdsperiode worden de notities terug bij de hand genomen en wordt nagegaan wat de leerlingen terecht hebben gebracht van de geformuleerde doelen.

Niet alleen het leren van leerlingen kan het onderwerp zijn van interviews. Ook andere aspecten van de manier waarop leerlingen het schoolgebeuren ervaren, kunnen aan bod komen. Hierbij denken we bijvoorbeeld aan studiemotivatie, inzet, doelmatigheidsbeleving en welbevinden.

5 Voorbeelden van leerlingenbetrokkenheid

In deze paragraaf presenteren we werkwijzen om leerlingen in de verschillende evaluatiefases te activeren. We richten ons eerst op een aantal algemene voorbeelden van leerlingenbetrokkenheid. Daarna ontwikkelen we materiaal dat bij de leerlingen uit de vooropgestelde casussen kan worden aangewend.

De mate van betrokkenheid van leerlingen moet worden opgevat als een continuüm dat bovendien kan variëren al naargelang de betreffende fase van het evaluatieproces. Leerlingen kunnen bijvoorbeeld sterk betrokken worden bij het zoeken naar criteria voor een bepaalde prestatie terwijl ze niet de kans krijgen om de criteria ook zelf effectief toe te passen. Dat neemt niet weg dat het veelal leerrijk is om leerlingen in de verschillende fases van het evaluatieproces te betrekken.

5.1 Algemene voorbeelden van leerlingenbetrokkenheid

Het betrekken van leerlingen betekent dat je als leerkracht tracht hen een rol te laten vervullen in allerlei taken die traditioneel enkel de verantwoordelijkheid waren van de leerkracht. Om een aantal algemene voorbeelden te geven van de betrokkenheid van leerlingen overlopen we in de volgende figuur nogmaals het evaluatieproces dat we in dit hoofdstuk voorstelden. Bij elke fase presenteren we een aantal werkwijzen die het mogelijk maken om de rol van leerlingen actiever in te vullen en om de verantwoordelijkheid van leerlingen te vergroten.

Figuur 4.43: Algemene voorbeelden van leerlingenbetrokkenheid

Stap 1: definiëren en exploreren van leerlingenbetrokkenheid

- ✓ Leerlingen kunnen bevraagd worden over de rol die ze willen spelen in hun eigen evaluatieproces. Zijn ze bereid om zichzelf en hun klasgenoten te beoordelen? Waarom (niet)? Welke functies zouden de op die manier verkregen evaluatiegegevens moeten hebben of net niet mogen hebben? Op welke manier denken ze dat ze kunnen leren uit een beoordeling door hun klasgenoten? Rekening houdend met de wensen en mogelijkheden van de leerlingen, kan op die manier een gepaste vorm van leerlingenbetrokkenheid op het getouw worden gezet.
- ✓ Leerlingen kunnen zelf kiezen wanneer ze bepaalde evaluatietaken willen afleggen. Of: in overleg met de leerkracht wordt een klassikaal evaluatiemoment afgesproken.
- ✓ Op welke manier moet de puntenverdeling in een co-evaluatie plaatsvinden. Voor welk systeem opteren de leerlingen? Waarom? Kunnen ze zelf een beter systeem ontwikkelen?
- ✓ Welke vorm van evaluatie is volgens de leerlingen het meest geschikt om te demonstreren of de vooropgestelde doelstellingen al dan niet bereikt zijn?

Stap 2: het bepalen van criteria

- ✓ De leerlingen krijgen de kans om zelf mee aan te geven welke criteria doorslaggevend zouden moeten zijn om het onderscheid te kunnen maken tussen goede en slechte prestaties. Zo kan je leerlingen bijvoorbeeld zelf op zoek laten gaan naar wat de kenmerken (criteria) zijn van een goed logboekverhaal. Dat kan door de leerlingen te vragen echte verslagen te beoordelen. De vraag hierbij is dan bijvoorbeeld: welke aspecten in de verslagen zijn belangrijk om een goed beeld te krijgen van wat zich heeft afgespeeld tijdens de les? Waar het ene verslag goed in is, kan het andere tekortschieten. Op die manier kunnen de leerlingen zelf essentiële kenmerken van de beoogde prestatie aan het licht brengen. Op een gelijkaardige manier kan men tewerk gaan om criteria voor een laswerk te beoordelen.
- ✓ De leerkracht geeft een lijst van mogelijke criteria en selecteert samen met de leerlingen welke criteria het meest gepast zijn om toekomstige leerlingenprestaties te beoordelen. Of men kan leerlingen bevragen over het gewicht dat elk van de criteria in de uiteindelijke beoordeling zou moeten krijgen.

Stap 3: het toepassen van de criteria

- ✓ Leerlingen gaan na in welke mate hun eigen prestaties aan de vooropgestelde criteria voldoen. Ze doen hierbij een beroep op één van de evaluatieinstrumenten die we voorstelden in dit hoofdstuk.

- ✓ Leerlingen doen evaluatieve uitspraken over een door de klasgenoten gemaakt product. Ze spreken aan de hand een holistische rubriek een algemene beoordeling uit. Deze beoordeling baseren ze op de resultaten van het invullen van een evaluatielijst.
- ✓ Nadat de leerkracht een groep leerlingen een algemeen groepscijfer heeft toegekend, bespreken de leden van de groep wie meer en wie minder punten verdient dan door de leerkracht aangegeven.

Stap 4: het geven van feedback

- ✓ De leerlingen beschrijven in hun logboek welke activiteiten ze volgens zichzelf goed uitgevoerd hebben en waarom dat zo is. Daarnaast geven ze aan op welke criteria ze nog tekortschieten en op welke manier ze van plan zijn hun eigen prestaties nog te verbeteren.
- ✓ Leerlingen verantwoorden waarom ze het al dan niet eens zijn met de beoordelingen die hun klasgenoten aan de hand van een evaluatielijst over hun prestaties uitspreken.
- ✓ Leerlingen geven aan waarom ze een prestatie van een bepaalde klasgenoot beter vonden dan de prestatie van een andere klasgenoot.
- ✓ Na een evaluatie door de leerkracht stellen de leerlingen een plan op dat de kwaliteit van de manier waarop ze in groep samenwerken zo groot mogelijk moet maken.
- ✓ De leerlingen vullen in naam van een klasgenoot een attitudeschaal in. Op basis van deze gegevens wordt in groepen van vier overlopen ten aanzien van welke aspecten leerlingen volgens hun medeleerlingen verrassend scoren. Waarom sluit de score op de Likert-schaal die de klasgenoot gaf (niet) aan bij de beoordeling die de leerling zichzelf zou geven?
- ✓ Op basis van eventuele tekorten in de eigen prestaties en de prestaties van klasgenoten, rapporteren leerlingen een aantal elementen die het mogelijk maken de door de leerkracht verzorgde instructie te verbeteren.

Stap 5: het stellen van doelen

- ✓ De leerlingen geven aan welke tussentijdse doelen ze nog moeten bereiken vooraleer ze de vooropgestelde vaardigheid volledig onder de knie hebben.
- ✓ De leerlingen geven voor een klasgenoot aan, aan welke aspecten van een complexe vaardigheid (een presentatie geven, een deurdial maken) hij in de komende weken extra aandacht zou moeten besteden.

5.2 Voorbeelden van leerlingenbetrokkenheid in onze casussen

Bij het illustreren van vaardigheidsevaluatie en authentieke evaluatie werden een aantal evaluatieinstrumenten voor schrijf- en luistervaardigheden Nederlands gepresenteerd. Het doel van deze instrumenten was dat ze door de leerkracht gebruikt zouden worden om de prestaties van leerlingen te evalueren. Om de leerlingen bij het beoordelingsproces te betrekken, kan men op een eenvoudige manier dezelfde instrumenten door de leerlingen zelf laten gebruiken. Men moet er dan uiteraard op toezien dat de manier waarop de criteria en standaarden beschreven worden, aangepast is aan kenmerken van de leerlingen uit de eerste graad.

Even stilstaan... 4.3: De geschiktheid van vroegere evaluatieinstrumenten

Welke van de rubrieken die we bij vaardigheidsevaluatie en authentieke evaluatie presenteerden, zijn volgens u reeds geschikt om door leerlingen gebruikt te worden? Wat zou u veranderen aan de rubrieken en beoordelingsinstrumenten die nog niet geschikt zijn?

5.2.1. Leerlingenbetrokkenheid bij leerlingen in de eerste graad - Nederlands

Leerlingen uit de eerste graad kunnen niet alleen zelf de evaluatieinstrumenten toepassen, ze kunnen ook betrokken worden bij het ontwikkelingsproces van criteria. Gesteld dat de evaluatieopdracht erin bestaat dat leerlingen in kleine groepjes een voorpagina van een krant moeten maken, dan kan dit door de leerlingen vooraf de volgende taak te geven.

Figuur 4.44: Leerlingen betrekken bij het ontwikkelen van criteria

Taak: bestuderen van authentieke voorpagina's. Leerlingen verzamelen elk twee voorbeelden van voorpagina's van echte dagbladen. Deze gaan ze eerst lezen, goed bestuderen en met elkaar vergelijken. Ze hebben hierbij zowel aandacht voor de vorm als de inhoud. Vervolgens beantwoorden de leerlingen drie open vragen, eventueel in kleine groep.

1. Wat valt je het meest op wanneer je de voorpagina's bekijkt?
2. Zie je grote verschillen tussen de twee voorpagina's?
 - ✓ Zo ja, welke zijn die verschillen?
 - ✓ Zo nee, waarin zijn de twee voorbeelden gelijk?
3. Waarom vind je dit een goede of een minder goede voorpagina?
 - ✓ Wat is er goed aan de voorpagina's?
 - ✓ Waarin schieten ze tekort?

Na het beantwoorden van deze vragen worden de antwoorden van de leerlingen klassikaal overlopen. Hierbij worden al de aangereikte en relevante elementen ondergebracht in twee groepen: (1) de elementen die bijdragen tot de kwaliteit van een voorpagina en (2) de elementen die maken dat er aan kwaliteit moet worden ingeboet. Eventueel wordt de inbreng van de leerlingen nog door de leerkracht aangevuld.

Indien alle elementen vervolgens op een positieve manier geformuleerd worden, bekomt men een lijst met potentiële criteria om de kwaliteit van een voorpagina te beoordelen. Die ziet er bijvoorbeeld als volgt uit: de lay-out trekt de aandacht naar de belangrijkste stukken; de titels beschrijven duidelijk de inhoud; verschillende berichten worden duidelijk van elkaar gescheiden; de naam van de krant is zeer duidelijk leesbaar; de inhoud is interessant en brengt nieuws; de lezer leert iets bij; er staan geen schrijffouten in; er worden gepaste grafieken en illustraties gegeven; alle artikels staan op zich (ze geven elk een antwoord op de vier W's aan: wie, wat, wanneer en waarom); er is informatie verzameld uit de eerste hand.

Op een dergelijke manier kunnen niet alleen criteria verzameld worden voor het succesvol samenstellen van een voorpagina. Dezelfde werkwijze kan ook bij andere tekstsoorten gebruikt worden. Door te kijken naar authentieke brieven, verhalen, presentaties, pleidooien, gedichten kunnen door de leerlingen telkens specifieke criteria naar voren worden geschoven. Uiteraard heb je hier als leerkracht een sturende functie. Niet relevante elementen moeten weggelaten

worden terwijl belangrijke kenmerken misschien meer aandacht verdienen dan door de leerlingen aangegeven wordt. Zo kan je leerlingen uitnodigen meer aandacht te besteden aan de structuur van een krantenbericht. Eventueel kan daartoe één van de analytische taalvaardigheidsrubrieken die we reeds presenteerden, worden gebruikt. Het komt er op aan de leerlingen te laten reflecteren over wat maakt dat een opiniestuk of krantenartikel al dan niet kwaliteitsvol is. Eenmaal ze weten waar het bij de verschillende criteria om draait, wordt het mogelijk die criteria aan te wenden bij het beoordelen van de eigen prestaties of de prestaties van klasgenoten.

De eigenlijke evaluatietaak van leerlingen bestaat erin zelf een voorpagina samen te stellen. Hierbij gaan de leerlingen in groep te werk. Het is aangewezen de leerlingen te stimuleren om het verloop van de taak systematisch te plannen en om de groepswerking voortdurend te bewaken. Hiertoe kunnen de leerlingen een logboek bijhouden. Het logboek kan bijvoorbeeld worden onderverdeeld in een groepsgeen en een individueel deel. Beide zijn sterk voorgestructureerd. De twee delen worden in de volgende figuren beschreven.

Figuur 4.45: Groepsgeen logboek

Doel van dit logboek is het plannen van de verschillende activiteiten. Jullie hebben elke dinsdag twee opeenvolgende lesuren die jullie kunnen besteden aan het maken van de voorpagina. De taak moet na vijf weken afgerond zijn. Aan het einde van de les vullen jullie gezamenlijk het volgende logboek in. Stop daarom tijdig met de andere activiteiten en voorzie zeker een kwartier.

Datum:	Lesuur nummer	Aanwezige leden van de groep:
	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	
Vandaag hebben we gewerkt aan... 1. Voltooid. 2. Begonnen.		
Volgende les gaan we verder met ... Op langere termijn staat het volgende nog op het programma.		
Hiertoe moeten we eerst het volgende thuis afwerken		
Taak	Verantwoordelijke	Tegen wanneer

Figuur 4.46: Individueel logboek

Naast een groepsgeen logboek, vullen jullie ook elke week een individueel logboek in. Hierin staan jullie eigen bedenkingen en ervaringen centraal. De nadruk ligt op de criteria van een goe-

de voorpagina en op de manier waarop jullie in groep samenwerken. Beantwoord elke dinsdagavond de volgende vragen. Bij het aanvangen van de volgende 2 lesuren overlopen jullie per groep de antwoorden van de week voordien.

1. Beantwoord de volgende vragen.
 - ✓ Op welke criteria scoort de voorpagina momenteel reeds voldoende goed?
 - ✓ Welke criteria moeten duidelijk nog worden bijgewerkt?
 - ✓ Wat was volgens jou de belangrijkste bijdrage die je de voorbije lesuren aan het werk hebt geleverd?
 - ✓ Op welke eigen activiteit of bijdrage ben je niet zo trots? Waarom niet?
2. Vul de evaluatielijst over 'samenwerken in groep' in met de voorbije twee lesuren in gedachten (zie figuur 4.47). Vul eventueel zelf nog criteria aan. Beantwoord daarna de volgende vragen.
 - ✓ Welke zijn de criteria waarop de groep volgens jouw nog tekort schiet (niet goed)?
 - ✓ Wat zou jullie groep moeten veranderen om hier beter op te scoren?

Figuur 4.47: Evaluatielijst 'samen in groep werken'

	Heel goed	OK	Niet goed	Niet van toepassing
Iedereen deed wat vooraf van hem verwacht werd.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Iedereen luistert naar de ideeën van anderen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We worden niet snel afgeleid.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Iedereen heeft zicht op onze toekomstige plannen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Het doel van de taken is voor iedereen duidelijk.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bij meningsverschillen vonden we een compromis.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Leerlingen worden in de loop van de taak steeds meer vertrouwd met de verschillende criteria. Dat betekent dat ook tijdens de eindevaluatie van hun product (de voorpagina) een beroep kan worden gedaan op de medewerking van de leerlingen. Leerlingen kunnen zowel hun eigen eindproduct als het eindproduct van andere groepen beoordelen. Ze kunnen hierbij hun klasgenoten van leerrijke en gerichte feedback voorzien.

5.2.2. Leerlingenbetrokkenheid bij leerlingen in de eerste graad wiskunde

Voor het vak wiskunde in de eerste graad van het secundair onderwijs willen we twee voorbeelden aanhalen. Eerst presenteren we een evaluatietaak die zowel rekenkundige vaardigheden evalueert als leerlingen uitnodigt tot reflectie. Daarna geven we een voorbeeld van een attitude-schaal die de houding van leerlingen ten aanzien van het vak wiskunde bevraagt.

1. Integratie van leren en evalueren

Eén van de eindtermen wiskunde voor de eerste graad luidt: leerlingen kunnen vanuit tabellen met cijfergegevens het rekenkundig gemiddelde en de mediaan berekenen en hieruit relevante informatie afleiden. Wanneer we leerlingen tevens willen aanzetten om over hun eigen prestaties na te denken, kunnen we hen deze rekentaken laten toepassen op eigen beoordelingscijfers. Dit gebeurt in het voorbeeld in figuur 4.48.

Figuur 4.48: Berekenen van gemiddelde en mediaan

Doorheen het voorbije semester hebben jullie vijf maal een herhalingsstoets gemaakt. Ik heb telkens de resultaten van heel de klas bewaard. Jouw eigen score en die van vier van jouw klasgenoten vind je in onderstaande tabel. De cijfers zijn echt maar de namen heb ik weggelaten. Bij deze tabel horen twee taken.

	jezelf	klasgenoot 1	klasgenoot 2	klasgenoot 3	klasgenoot 4
Toets 1 (op 10 punten)	6	8	5	9	6
Toets 2 (op 20 punten)	13	17	11	20	12
Toets 3 (op 20 punten)	13	16	10	16	13
Toets 4 (op 10 punten)	7	8	4	9	7
Toets 5 (op 10 punten)	7	9	4	8	7
Gemiddelde					
Mediaan					

Taak 1. Bereken voor jezelf en je vier klasgenoten de gemiddelde score en de mediaan. Geef bij je antwoord alle bewerkingen weer die je hierbij uitvoert. Beantwoord daarna de volgende vragen.

- ✓ Wie heeft over de vijf toetsen samen het beste resultaat behaald?
- ✓ Is er een groot verschil tussen de prestaties van de vijf leerlingen?

Taak 2. Stel jezelf even in de plaats van de leerkracht en geef zelf commentaar bij jouw eigen prestaties. Geef in jouw commentaar minstens één antwoord op de volgende vragen (bijkomende informatie mag uiteraard ook worden opgenomen).

- ✓ Ben je tevreden over je eigen prestaties?
- ✓ Hoe goed presteerde je in vergelijking met de rest van de groep?
- ✓ Waarom denk je dit dat zo is? Of: hoe verklaar je dat?
- ✓ Wat zou je jezelf adviseren om nog betere wiskunderesultaten te bereiken?

Om de antwoorden van de leerlingen op een dergelijke taak te beoordelen, moet er een duidelijk onderscheid worden gemaakt tussen het rekenwerk en de reflectie. Voor het rekenwerk gelden dezelfde criteria als bij voorheen gepresenteerde rekentaken. In feite is deze rekentaak immers een voorbeeld van een authentieke evaluatietaak. Ook voor de reflectie van de leerling kunnen criteria vooropgesteld worden. We kiezen ervoor een beoordelingsinstrument te presenteren dat bij verschillende reflecties gebruikt kan worden. Hiertoe presenteren we eerst een evaluatielijst. In figuur 4.50 geven we daarna een holistische rubriek weer die het mogelijk maakt algemene beoordelingen uit te spreken.

Figuur 4.49: Evaluatielijst voor reflectie

	1	2	3	NVT
Erkennen van sterke en zwakke punten.	0	0	0	0
Verklaring geven voor succes en mislukking.	0	0	0	0
Aangeven wat er geleerd werd.	0	0	0	0
Persoonlijke ervaringen en gevoelens aan bod brengen.	0	0	0	0
Toekomstige doelen vooropstellen.	0	0	0	0
Aangeven hoe die doelen bereikt zullen worden (strategie).	0	0	0	0
De reflectie is duidelijk (een vreemde kan het geheel begrijpen).	0	0	0	0
Hoofd- en bijzaken onderscheiden (enkel relevante zaken komen aan bod).	0	0	0	0
Er is een kritische houding aannemen t.o.v. zichzelf.	0	0	0	0
1=goed, 2=volgende met leemtes, 3=onvoldoende, NVT=niet van toepassing				

Figuur 4.50: Holistische rubriek 'Goed reflecteren'

5 - Zeer goed. De reflectie is zeer duidelijk, accuraat en gedetailleerd. De leerling omschrijft goed waarover hij/zij reflecteert en wat er geleerd werd. De leerling heeft een kritische houding ten opzichte van zichzelf. Zowel goede als minder goede elementen komen aan bod. Waar mogelijk wordt een verklaring gegeven voor de prestaties. De leerling stelt voor zichzelf duidelijke en haalbare doelen voorop. Daarbij wordt goed aangegeven op welke manier deze doelen bereikt zullen worden. Naast de nodige feiten en vaststellingen brengt de leerling op een gepaste manier ook persoonlijke ervaringen en gevoelens aan bod.

4 - Goed.

3 - Voldoende met leemtes. De reflectie is voldoende duidelijk en accuraat. De inhoud van de reflectie en het leerproces zijn verstaanbaar. De leerling neemt een kritische houding aan maar zonder veel diepgang. Er wordt een aanzet gegeven om te verklaren waarom de prestaties zijn wat ze zijn. Er zijn doelen maar die zijn niet allemaal duidelijk of haalbaar. Er is een strategie. Persoonlijke ervaringen en gevoelens komen aan bod.

2 - Onvoldoende.

1 - Slecht. De reflectie is onduidelijk en bevat niet-relevante elementen. Het is onduidelijk waarover de leerling reflecteert. Er wordt geen kritische houding aangenomen ten opzichte van zichzelf. Er worden geen verklaringen gegeven. Doelen en een strategie om die te bereiken, zijn afwezig. Persoonlijke ervaringen en gevoelens worden niet opgenomen.

2. Attitudes bij het vak wiskunde

Attitudes hebben vaak een belangrijke invloed op de prestaties van leerlingen. Dit geldt ook voor het vak wiskunde. Zowel voor de leerkracht als voor de leerlingen is het daarom van belang om regelmatig stil te staan bij de houding van leerlingen ten opzichte van het vak. Uit toelichting bij de eindtermen voor de eerste graad blijkt dat attitudes belangrijke onderwijsdoelen zijn en dus ook geëvalueerd moeten worden (DVO, 1997).

Figuur 4.51: Attitudes bij wiskunde in de eerste graad secundair onderwijs

De motivatie van de leerling kan verhoogd worden door er voor te zorgen dat het vak nuttig, zinvol en boeiend ervaren wordt. Het boeiende wordt bereikt als de leerlingen in bewondering kunnen staan voor de schoonheid en de perfectie van een meetkundige figuur, de helderheid van een redenering en de elegantie van een formule. Als de leerlingen ontdekken dat ze bekwaam zijn om hun groeiende wiskundekennis te gebruiken in nieuwe situaties, groeit hun vertrouwen en worden ze zelfzekerder. De leerlingen moeten ervaren dat wiskunde praktisch nut heeft, dat ze een vormende en esthetische waarde heeft. Leerlingen hebben de goede gewoonte kritisch te staan tegenover allerlei cijfermateriaal, tabellen, berekeningen, enz. Ze zijn bereid een probleem zelf aan te pakken. Het leren door vallen en opstaan moet niet ontmoedigend werken. Uit fouten en verkeerde keuzes kan eveneens geleerd worden. De leerlingen verwerven de goede gewoonte op hun oplossingsproces terug te kijken en hun resultaat te toetsen. Ze ervaren het oplossingsproces even waardevol als het resultaat.

De houding die leerlingen hebben ten opzichte van wiskunde wordt in de realiteit zelden geëvalueerd. Toch zijn er hiertoe goede instrumenten beschikbaar. We verwijzen hier naar een attitudeschaal van Harskamp (1988). Deze schaal bevat verschillende items die zich richten op de positieve/negatieve houding van leerlingen tegenover wiskunde. Het gaat erom hoe boeiend, interessant of vervelend de leerling de lessen wiskunde vindt. Daarbij komen een aantal zogenaamde subschalen aan bod: gevoelens, gedragingen, verwachtingen, wensen en motieven. De score van de leerling op elk van die subschalen wordt berekend door het gemiddelde te nemen van de verschillende items in de attitudeschaal die betrekking hebben op deze subschaal. De schaal en de items worden in figuur 4.52 weergegeven.

Figuur 4.52: Attitudeschaal 'houding tegenover wiskunde'

De instructies. In deze vragenlijst staan een aantal beweringen over het vak wiskunde. We willen graag weten of je het daarmee eens of oneens bent. Je moet daartoe kiezen tussen één van de volgende mogelijkheden:

- ✓ (4) helemaal mee eens
- ✓ (3) beetje mee eens
- ✓ (2) beetje mee oneens
- ✓ (1) helemaal mee oneens

Maak zoveel mogelijk gebruik van alle vier de antwoordmogelijkheden, maar maak van je keuze geen al te groot probleem. Je antwoord moet natuurlijk wel zo goed mogelijk je eigen mening weergeven.

De verschillende items zijn:

- ✓ Wiskunde zal niet gauw een hobby van mij worden.
- ✓ Ik vind het vervelend om rekenopgaven te maken.
- ✓ Onze wiskundelessen zijn vaak boeiend en interessant.
- ✓ In het latere leven kun je best zonder de wiskunde die je op school leert.
- ✓ Ik maak wel eens meer wiskundeopdrachten dan we opgekregen hebben.
- ✓ Na schooltijd doe ik wel eens spelletjes die met wiskunde te maken hebben.
- ✓ Bij veel dingen die je iedere dag tegenkomt, heb je veel aan wat je in de wiskundelessen leert.
- ✓ Buiten de school heb je weinig aan wat je in de wiskundeles leert.

- ✓ Ik geloof dat de wiskundelessen maar weinig nut hebben.
- ✓ Ik heb weinig zin om me na schooltijd ook nog bezig te houden met dingen uit de wiskundeles.
- ✓ Wiskunde is van belang om goed met geld te kunnen omgaan.
- ✓ Ik wil best meer lessen wiskunde in de week hebben.
- ✓ Ik werk wel eens een stukje vooruit in ons wiskundeboek.
- ✓ Ik vind de wiskundelessen saai.
- ✓ Wiskunde is van belang om later een baan te krijgen.
- ✓ Ik merk aan andere vakken dat ik wat heb aan de wiskundelessen.
- ✓ Ik wil in mijn vrije tijd wel eens een paar rekenpuzzels oplossen.
- ✓ In de wiskundeles gaat de tijd snel voorbij.
- ✓ Ik denk dat je later op school weinig hebt aan de wiskundelessen die we nu krijgen.

Naast de houding ten aanzien van het vak zijn er nog verscheidene niet-vakspecifieke attitudes van belang. We denken hierbij bijvoorbeeld aan leergierigheid, inzet, doorzettingsvermogen, zin voor initiatief, efficiëntie, orde en nauwkeurigheid. Voor elk van deze attitudes kunnen rubrieken worden opgesteld. Die kunnen zowel gebruikt worden om beoordelingen uit te spreken over het gedrag van leerlingen als om leerlingen een basis te geven om over hun eigen prestaties te reflecteren. Dat blijkt uit de rubriek in figuur 4.53 die de attitudes van leerlingen ten aanzien van werkmethode en efficiëntie bij het oplossen van wiskunde problemen onder de loep neemt (gebaseerd op Soudant, 2002). Met minimale aanpassingen kan de rubriek echter ook gebruikt worden om de aanpak van leerlingen bij andere inhouden te evalueren (een deurskader maken, een elektrisch probleem opsporen of een Latijnse vertaling maken).

Figuur 4.53: rubriek voor de attitude 'efficiëntie en zoeken naar oplossingsmethoden'

(5) **Uitstekend.** De leerling heeft een efficiënte zelfcontrole en stuurt de werkmethode zelf bij (in overleg met de leraar). In moeilijke situaties gaat hij nog efficiënter werken zonder zich van de wijs te laten brengen. De leerling vindt creatieve oplossingen.

(4) **Goed.** De leerling begint niet aan verschillende taken tegelijk, werkt zijn taken af en heeft een goede werkplanning. De leerling kiest voor praktische oplossingen en neemt efficiënte besluiten. Nu en dan vindt hij een originele en creatieve oplossing. De leerling staat open voor vernieuwing en kan zich aan nieuwe methodieken aanpassen.

(3) **Matig - voldoende.** De leerling doet pogingen om planmatig te werken en om een goede werkvoorbereiding te maken. Hij volgt zelfstandig eenvoudige werkrouines en standaardprocedures stap voor stap op. Voor het oplossen van problemen kan de leerling enkel kiezen tussen bestaande alternatieven.

(2) **Onvoldoende.** Alleen wanneer de leerling ondersteuning krijgt, wordt een goede werkmethode gevolgd. Het probleem wordt nogal impulsief, zonder planning aangevat. De leerling verzint oplossingen zonder ze op hun efficiëntie te toetsen. Het tijdsgebruik is chaotisch en er is geen werkmethode.

(1) **Rampzalig.** De leerling heeft lak aan methodieken en procedures. Hij gooit er met de pet naar.

5.2.3. Leerlingenbetrokkenheid bij leerlingen Elektriciteit/Elektronica

Bij de vaardigheidstaak voor de studierichting Elektriciteit/Elektronica presenteerden we een checklist om te evalueren of leerlingen succesvol de verschillende stappen in het opsporen van fouten in een elektrische kring beheersen (zie figuur 4.14). Deze checklist kan zowel door de leerkracht als door de leerling zelf of diens medeleerlingen gehanteerd worden. De rol van de leerlingen in de evaluatie van de gepresenteerde taak kan dus nog uitgebreid worden. De verschillende vragen in de checklist nodigen leerlingen uit om over hun eigen aanpak te reflecteren en ze geven -zo nodig- indicaties voor een betere aanpak. Op een gelijkaardige manier kunnen (mede)leerlingen betrokken worden in het evalueren van allerlei andere doelstellingen binnen de richting Elektriciteit/ Elektronica (zoals gebruik maken van meettoestellen, ontbramen van elektriciteitsbuizen, verslagen maken).

Eén van de doelstellingen bij leerlingen Elektronica die het noodzakelijk maken om de leerlingen te activeren in het leer- en evaluatieproces is hun houding ten aanzien van het toepassen van veiligheidsvoorschriften. In de tweede graad geldt dit bijvoorbeeld bij het gebruik van meetinstrumenten. Aan de hand van zelfevaluatielijsten zoals de volgende kunnen leerlingen hun eigen gedragingen en houding onder de loep nemen.

Figuur 4.54: Zelfevaluatielijst 'toepassen van veiligheidsvoorschriften'

1 - Zeker en vast – Altijd	Mijn beoordeling
2 - Ja, doorgaans wel	4 3 2 1
3 - Durf ik al eens vergeten	
4 - Hier schiet ik echt tekort	
Ik ben mij bewust van de gevaren van elektrocutie.	
Vooraleer een voltmeter aan te sluiten, besteed ik aandacht aan de spanningssoort en de grootte van de spanning.	
Ik ben er mij van bewust dat bij het meten van stroom en spanning de installatie, de meetsnoeren en het meettoestel onder spanning staan.	
Vooraleer de ampèremeter aan te sluiten, besteed ik aandacht aan de stroomsoort en de grootte van de stroom.	
Ik ben mij bewust van de gevaarlijke invloed van vochtigheid op de geleidbaarheid van elektrische stroom.	
Na het gebruik van elektrische instrumenten trek ik het stopcontact uit.	
Ik trek niet aan het snoer maar aan het stopcontact bij het uittrekken van kabels.	
Ik gebruik altijd geïsoleerd gereedschap.	
Ik vervang onmiddellijk defecte schakelaars, stekkers, e.d.	
Wanneer ik twijfel over de juiste aanpak van een probleem, vraag ik hulp.	
Ik heb aandacht voor pictogrammen die wijzen op gevaar.	
Ik kijk steeds na of ik wel goedgekeurd installatiemateriaal gebruik.	
Ik ga <u>niet</u> roekeloos te werk.	
Totaal: gemiddelde score...	

De vragenlijst leidt zowel tot een algemene score als tot gedetailleerde informatie over concrete aspecten betreffende de veiligheidsvoorschriften. De algemene score is echter enkel indicatief. Gemiddelde scores kunnen immers interessante informatie verdoezelen. Een leerling die, op twee keer 1 na, telkens 4 scoort, heeft een gemiddelde van 3,5. Dit kan de indruk wekken dat er

niets aan de hand is terwijl hij toch op twee veiligheidsvoorschriften duidelijk tekort schiet. Het is dus vooral de combinatie van de totale score en de afzonderlijke items die tot leerrijke feedback kan leiden.

5.2.4. Leerlingenbetrokkenheid bij leerlingen Verzorging

De stages die leerlingen uit de derde graad Verzorging lopen, bieden interessante mogelijkheden om hen te betrekken bij hun eigen evaluatie en om hen uit te nodigen om te reflecteren over hun eigen leerproces. De stageplaatsen zijn echter zeer verscheiden. Leerlingen kunnen terecht in kinderdagverblijven, peutertuinen, onthaalgezinnen, bejaardentehuizen, thuiszorg, instellingen voor gehandicapten, etc. Om in elk geval een kwaliteitsvolle evaluatie van de stage mogelijk te maken, wordt er best per type stageplaats een evaluatieformulier voorbereid. Inzake kennis, attitudes en vaardigheden zullen de doelstellingen die via een leerlingenstage verwezenlijkt moeten worden, er immers telkens anders uitzien. De evaluatie van een stage is dus niet eenvoudig. Vandaar dat die onze aandacht verdient.

Doorgaans bevat een evaluatiedossier van een leerling verschillende elementen: de evaluatieverslagen van de stagementor (de begeleider in de stageplaats), de verslagen van de stagebegeleider (de leerkracht) en het stageschrift van de leerling. In de volgende paragrafen rapporteren we een aantal mogelijkheden om leerlingen in hun eigen stageschrift te activeren. We richten ons hierbij op twee belangrijke doelen van een dergelijk stageschrift: (1) het aanleggen van een verslag van de stageactiviteiten en (2) de zelfevaluatie van de leerling. Beide concretiseren we aan de hand van een stage in een rusthuis.

1. Rapporteren van stageactiviteiten en -ervaringen

Bij het omschrijven van logboeken stelden we dat er - wat inhoud van deze zelf-evaluatieinstrumenten betreft - drie groepen worden onderscheiden. Deze drie inhoudelijke verschillen vormen telkens een andere manier om naar de stageactiviteiten te kijken. We geven van elk een voorbeeld.

a. De taken van leerlingen

Een deel van het stageschrift kan zich toelagen op de dagelijks activiteiten van leerlingen. Op basis van vooraf in de stageformulieren vastgelegde activiteiten, krijgen de leerling en diens begeleiders een goed overzicht van welke taken reeds uitgevoerd werden. Tegelijkertijd wordt duidelijk welke activiteit de leerling nog niet of onvoldoende heeft ondernomen. Aan deze taken kan dan in de toekomst bijkomende aandacht worden besteed. Een dergelijke activiteitenlijst ziet er bijvoorbeeld als volgt uit. Een deel van de activiteiten is opgegeven terwijl leerlingen ook de kans hebben om zelf activiteiten toe te voegen.

Figuur 4.55: Activiteitenlogboek

Activiteiten	Week van tot	Ma	Di	Wo	Do	Vrij
--------------	--------------------------	----	----	----	----	------

Bad geven	x	x		x	x
Toilet aan lavabo		x	x	x	x
Toedienen van medicatie					x
Reinigen van toiletstoel		x	x		
Inrollen van haar					
Helpen met eten	x	x	x	x	x
.....					
Dansen met bejaarden			x		
Naar de markt gaan				x	

b. Het leren van leerlingen - learner-report

In hun stageschrift moeten leerlingen ook hun eigen leerervaringen onder woorden brengen. De stagiaire in het rusthuis beantwoordt daartoe bijvoorbeeld de volgende vragen.

Figuur 4.56: Learner report - leerlogboek

Op het einde van een stageweek sta je stil bij de activiteiten die je de voorbije week hebt ondernomen en de gesprekken die je op je stageplaats hebt gevoerd. Beantwoord daartoe telkens volgende vragen.

- Deze week heb ik nieuw geleerd dat...
- Ik ben voornamelijk beter geworden in (het)...
- Dat ik mezelf nog moet verbeteren, bleek deze week uit...
- Om beter te worden op deze punten, ben ik van plan om...
- Waarover zou je meer willen leren...
- Welke beoordeling verdien je volgens jezelf voor deze stageperiode? Waarom?

c. De ervaringen en gevoelens van leerlingen

Naast het plannen van activiteiten en het reflecteren over het eigen leren moeten de leerlingen ook de kans krijgen persoonlijke ervaringen en gevoelens te rapporteren. In het persoonlijk logboek bepaalt de leerling in grote lijnen zelf over welke onderwerpen hij schrijft. Dit neemt niet weg dat de leerkracht voor een zekere sturing kan zorgen. Dit blijkt bijvoorbeeld uit de richtvragen voor de leerlingen Verzorging. Er wordt een onderscheid gemaakt tussen aanvulzinnen, open vragen en het reageren op stellingen.

Figuur 4.57: Richtvragen voor een persoonlijk logboek

- ✓ Ik ben trots op...
 - ✓ Ik heb me slecht gevoeld toen...
 - ✓ Wat me van deze week zeker zal bijblijven is dat...
 - ✓ Ben je tevreden over deze stageweek? Waarom (niet)?
 - ✓ Zou je later graag in deze stageplaats werken? Waarom (niet)?
 - ✓ Waarom ga je al dan niet akkoord met volgende stellingen?
- Deze stageperiode mag voor mij gerust een week langer duren.
De stagementor heeft een correct beeld van mij.
Deze stageweek is echt tijdverlies geweest.

Het onderscheid tussen de verschillende inhoud van het stagelogboek bij de leerlingen Verzorging werd enkel ter illustratie aangehouden. In de realiteit richten logboeken en de reflecties van leerlingen zich veelal op een combinatie van deze verschillende inhoud.

2. Zelfevaluatie bij de stageactiviteiten

Om leerlingen te betrekken bij het evalueren van hun eigen stageactiviteiten wordt vaak een beroep gedaan op evaluatielijsten. Feit is dat om de stageprestaties van leerlingen te beoordelen sowieso criteria gebruikt worden. Ook als leerlingen niet betrokken worden bij de evaluatie moet je beschikken over een goede batterij criteria. Wanneer die dan toch ter beschikking zijn, vergt het slechts beperkte inspanningen om ze ook door de leerlingen zelf te laten gebruiken. In het hoofdstuk van de authentieke evaluatie presenteerden we reeds een evaluatielijst voor de stage-evaluatie in een kleuterschool. Gelijkaardige evaluatielijsten bestaan om het stageverloop in een bejaardentehuis te evalueren (zie fig. 4.58). Deze lijsten kunnen zowel door de stagebegeleider als door de leerling zelf ingevuld worden. Of: waarom niet -indien mogelijk- de bejaarden zelf vragen om het evaluatie-instrument (gedeeltelijk) in te vullen.

Figuur 4.58: Evaluatielijst 'stage bejaardentehuis'

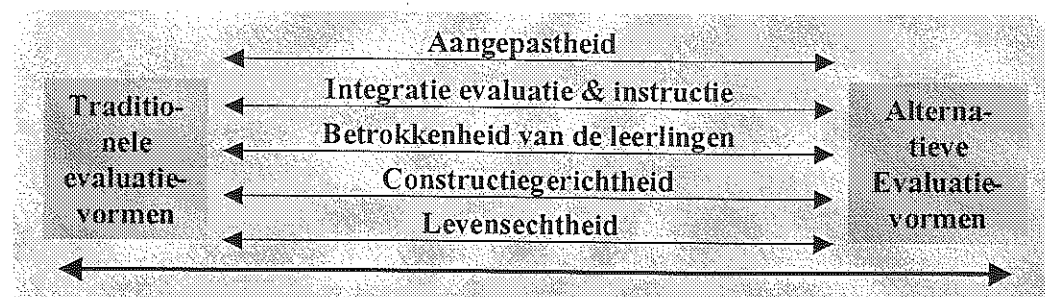
Evaluatiecriteria	Week 1		Week 2		Aandachtspunten voor het verdere verloop van de stage
	LL	SB	LL	SB	
Zich aan de afspraken houden					
Polsslag opnemen en aantekenen					
Verzorgingsdossier invullen					
Zorgzaam omgaan met apparatuur					
Bedbad toedienen					
...					
Algemene beoordeling					
(LL - leerling; SB - stagebegeleider)					
Beoordelingsschaal:					
✓ (+) Doelstelling bereikt					
✓ (±) Doelstelling gedeeltelijk bereikt, nog bij te sturen					
✓ (-) Doelstelling niet bereikt (motiveer telkens waarom)					
✓ (nvt) niet van toepassing (niet geobserveerd, heeft zich niet voorgedaan, ...)					

Zich baserend op de gegevens uit de evaluatielijst kunnen leerlingen zichzelf een omvattende beoordeling geven. Dat kan bijvoorbeeld aan de hand van een gelijkaardige holistische rubriek als diegene die gepresenteerd werd bij het illustreren van authentieke evaluatie bij leerlingen Verzorging (zie fig. 4.28). Het gebruik van dergelijke (zelf)evaluatielijsten en rubrieken houdt uiteraard niet op bij het invullen ervan. Op basis van de verkregen gegevens en eventuele verschillen tussen de beoordeling van de stagebegeleider en de leerlingen zelf, biedt de lijst immers tal van reflectiemogelijkheden en leerkansen.

Hoofdstuk 4: Het gebruik van portfolio's

De term 'portfolio' klinkt wellicht bekend in de oren. Verschillende professies hebben in het verleden reeds gebruik gemaakt van portfolio's om competenties, vaardigheden en verwezenlijkingen aan te tonen. Te denken valt onder andere aan schilders, modellen, journalisten en architecten. Pas recent is men portfolio's gaan hanteren als een onderwijskundige evaluatiemethode. In die context is een portfolio meer dan een uitstallast voor wat men in zijn mars heeft of reeds bereikt heeft. De onderwijsportfolio verschilt in een aantal opzichten van de traditionele artistieke portfolio. In de eerste plaats komt het er niet (alleen) op aan jezelf zo goed mogelijk voor te stellen. De nadruk ligt ook op de groei van de leerling. Men richt zich op de brede ontwikkeling van de leerling en niet zozeer op specialisaties. Bovendien bevat een onderwijskundig portfolio de reflectie van de leerling op de eigen werkstukken (Moya & O'Malley, 1994).

Uit deze kenmerken blijkt reeds dat de portfolio een interessante evaluatiemethode kan zijn. Meer nog: wanneer de portfolio in de juiste context gebruikt wordt en gebaseerd is op de juiste taken, kan deze evaluatievorm tegemoet komen aan elk van de kenmerken van alternatieve evaluatievormen (zie fig. 4.59). Vandaar de interesse voor het gebruik ervan.



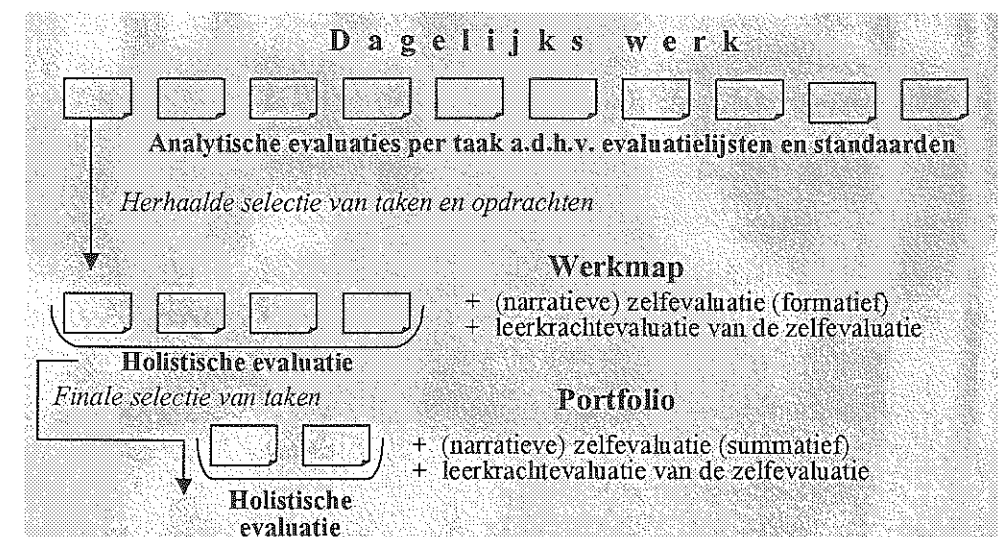
Figuur 4.59: De alternatieve kenmerken van portfoliogebruik

1 Portfolio geïllustreerd

In volgende illustratie en in de verdere beschrijving van portfolio's is er herhaaldelijk sprake van opdrachten en werkstukken. Deze vormen dan ook de basis van de portfolio. Met opdrachten en werkstukken bedoelen we de verscheidenheid aan antwoorden en producten die leerlingen tijdens het dagelijkse werk op school of thuis maken. Ze bestaan bijvoorbeeld uit verschillende grafieken, verschillende oplossingen van vergelijkingen, verschillende opstellen en vertalingen of verschillende lasverbindingen. In wezen kan elke leerkracht bij het volgende verloop van een portfolio eigen werkstukken in het achterhoofd houden.

Het gebruik van portfolio heeft geen vastliggende structuur, maar wel vastliggende ingrediënten. Volgend voorbeeld is dan ook maar één mogelijke invulling. Het verloop van devoorbeeldportfolio wordt in figuur 4.60 schematisch weergegeven (gebaseerd op ASCD, 1996). Uit deze figuur blijkt dat in het tot stand komen van deze portfolio drie stappen kunnen onderscheiden worden.

Figuur 4.60: Een mogelijk gebruik van een portfolio



In de eerste fase van de evaluatie verzamelen de leerlingen alles waaraan ze op de één of andere manier gewerkt hebben. Dit betekent niet meer dan dat het dagelijkse werk systematisch wordt bijgehouden. Zowel de taken als de evaluaties van het dagelijkse werk komen zoveel mogelijk tegemoet aan de eisen die we hoger vooropstelden om van goede (alternatieve) evaluatie- en leertaken te kunnen spreken. In dit voorbeeld gebruikt de leerkracht (authentieke) vaardigheidstaken die aan de hand van analytische rubrieken geëvalueerd worden. Dit betekent dat leerlingen per taak op elk van de criteria die voorop gesteld werden feedback krijgen. Er wordt doorheen de lesperiode duidelijk bijgeleerd. De kwaliteit van de verschillende beschikbare taken is dus verschillend.

Afhankelijk van de doelstellingen van de portfolio en van leerinhoud die we willen evalueren, worden in een volgende stap een aantal van de dagelijkse opdrachten en taken geselecteerd en samengebracht in een werkmap. Deze werkmap is een voorloper van de eigenlijke portfolio. We noemen de werkmap ook wel formatieve of tussentijdse portfolio. Het doel is het leerproces van de leerlingen bij te sturen. De leerlingen selecteren taken om in hun werkmap te steken. Ze verantwoorden uitvoerig waarom ze net die taken opnemen. Het belang van reflectie bij het gebruik van portfolio's zal nog herhaaldelijk benadrukt worden. In ons voorbeeld is het zo dat er zonder zelfevaluatie van de leerling én evaluatie door de leerkracht geen sprake is van een volwaardige werkmap. Over de taken in de werkmap worden algemene, holistische beoordelingen uitgesproken. Belangrijk is tevens het feit dat de werkmap inhoudelijk nog kan evolueren. Naarmate het

aantal taken waaruit kan geselecteerd worden groter wordt, moeten de leerlingen de inhoud van en de reflecties over de werkmap verder aanpassen.

Pas wanneer we een eindbeoordeling willen uitspreken over de leerling gaan we over tot de volgende stap. Wanneer de portfolio gehanteerd wordt als summatief evaluatie-instrument komt het erop aan op het geschikte tijdstip uit de werkmap een finale selectie van documenten te maken. Samen met de zelfevaluatie en de evaluatie door de leerkracht ontstaat op die manier de summatieve portfolio.

2 Portfolio gedefinieerd

De term 'portfolio' is vrij algemeen en kan verschillende doelen, gebruiken, inhouden en vormen omvatten. Arter & Spandel (1992) en ASCD (1996) geven de volgende definitie van portfolio's in een onderwijscontext.

Een portfolio is een doelgerichte verzameling van het werk van een leerling waarin diens inspanningen, vooruitgang en verwezenlijkingen aangetoond worden. Deze verzameling moet de deelname van de leerlingen in het selecteren van de inhoud van de portfolio, in het selecteren van criteria voor selectie van de inhoud en in het selecteren van criteria ter beoordeling van hun werk omvatten. De portfolio omvat tevens zowel duidelijke bewijzen van de reflectie van de leerling over diens eigen prestaties als van de bemerkingen van de leerkracht op het werk en de reflectie van de leerling.

Het gebruik van een portfolio houdt in dat de leerling een verzamelmap aanlegt met informatie over zijn persoonlijke prestaties en ervaringen. Het gebruik van portfolio's in een educatieve omgeving houdt dus in dat de leerling beschikt over een bundel werkstukken die een beeld schept van zijn mogelijkheden en groei. Essentieel is dat de leerling zelf een selectie maakt uit de producten die hij doorheen een bepaalde tijdsperiode creëerde en dat deze selectie verantwoord wordt. De betrokkenheid van leerlingen wordt sterk benadrukt. Dit blijkt niet alleen uit het ontwikkelen en gebruiken van evaluatiecriteria, maar ook uit de systematische reflectie op het eigen leerproces.

Het gebruik van portfolio's kent verscheidene toepassingen. De evaluatievorm kan binnen verschillende onderwijsvakken afzonderlijk aangewend worden of zich interdisciplinair richten op een combinatie van vakken. En ook al is er in deze definitie sprake van leerlingen, toch kan men portfolio's hanteren op verschillende onderwijsniveaus, gaande van het kleuteronderwijs tot hoger onderwijs. Verscheidenheid in het gebruik van portfolio's kan ook betrekking hebben op de doelgroep waaraan de portfolio zal voorgesteld worden. Te denken valt aan het betrekken van ouders, van de directie van de school, van leerkrachten van het volgende schooljaar of de toekomstige werkgever en geïnteresseerde derden. Portfolio's kunnen bovendien ook worden aangewend om kwaliteitsaspecten van klassen, scholen en scholengemeenschappen te evalue-

ren. In deze beschrijving zullen wij ons echter toelagen op het gebruik ter evaluatie van individuele leerlingen.

Om het definiëren van de term portfolio af te ronden, willen we benadrukken dat de term ook vaak foutief gebruikt of misbruikt wordt. Hiermee doelen we bijvoorbeeld op evaluatiesituaties waarin van leerlingen of studenten verwacht wordt dat ze alle taken en opdrachten die ze volbrengen, samenbrengen in een handige map. Deze informatie wordt dan door de leerkracht gebruikt om beoordelingen uit te spreken. In dergelijke gevallen is er geen sprake van selectie van werkstukken, van verantwoording en van reflectie. Goede portfolio's zijn dus duidelijk meer dan een praktische verzameling van het gedane werk.

3 Kenmerken en doelen van een portfolio assessment

Ook al bestaat er geen éénduidige en enige correcte manier om een portfolio als evaluatiemethode te gebruiken, essentieel is alvast wel dat leerlingen verwacht worden gegevens te selecteren en te verzamelen, en over deze gegevens te reflecteren. In volgende alinea's duiden we een aantal belangrijke kenmerken van portfolio's. In paragraaf 3.2 staan we vervolgens stil bij mogelijke doelen van een portfolio.

3.1 Kenmerken van een portfolio assessment

Een portfolio-evaluatie is een proces dat verscheidene facetten kent. Volgende alinea's geven een overzicht van de belangrijkste kenmerken. Deze zijn selectie, verantwoording, reflectie en dialoog (Office of Research, 1993; George, 1995).

a. Selectie

Een portfolio bestaat niet uit de verzameling van alle elementen die de leerling in de studieperiode produceerde, noch uit een toevallige selectie uit die elementen. De inhoud moet zorgvuldig geselecteerd worden. Die selectie moet aantonen dat er effectief iets geleerd werd en dat de leerling effectief iets kan. Kortom: wat diens verwezenlijkingen en vooruitgang zijn (Birenbaum, 1995). De inhoud van de portfolio moet vanzelfsprekend aansluiten bij het onderwezen curriculum en moet dus een beeld geven van al de relevante aspecten van het leerproces van de leerling. Kenmerkend voor portfolio's is verder dat de inhoud over een bepaalde tijdsperiode kan evolueren. Op die manier kan de ontwikkeling van de maturiteit en van het inzicht van de leerlingen in het onderwerp en in het eigen leren duidelijk worden. In paragraaf 6 gaan we dieper in op de mogelijke inhoud van een portfolio.

b. Verantwoording en reflectie

Een essentieel kenmerk van het gebruik van portfolio's is de noodzaak om de zelfreflectie van de leerling op te nemen. Leerlingen moeten reflecteren over elk element dat

ze in de portfolio opnemen. Leidende vragen hierbij zijn waarom ze dat specifieke element hebben opgenomen, hoe ze het gecreëerd hebben, welke de moeilijkheden waren, wat ze leerden van de taak etc. Net door te moeten verantwoorden waarom ze bepaalde werkstukken al dan niet in de portfolio hebben opgenomen, worden leerlingen gestimuleerd om te reflecteren over hun eigen leerproces. Op het moment dat de portfolio ter evaluatie voorgelegd zal worden, verwacht men van leerlingen tevens dat ze een begeleidende uitleg schrijven die een weergave is van deze reflectie en die, aangepast aan de doelgroep waaraan de portfolio wordt voorgesteld, het achterliggende principe van de inhoud van de portfolio toelicht.

c. Dialoog

Een portfolio is dus een voortdurend proces dat zowel formatieve als summatieve kansen biedt om de vooruitgang van leerlingen in kaart te brengen. Het bevordert de dialoog en de samenwerking tussen de leerling en de leerkracht door de verwezenlijkingen en de ontwikkeling van de leerling gezamenlijk en systematisch te evalueren. Op die manier krijgt men een gevarieerd beeld van de prestaties en het leerproces van de leerling en krijgt men zicht op de sterktes en zwaktes van de betreffende leerling. Een portfolio vertelt als het ware het verhaal van de leerling als lerende.

Om van een portfolio als alternatief evaluatie-instrument te kunnen spreken moeten leerlingen zelf een selectie maken uit de beschikbare werkstukken om hun eigen capaciteiten of groei aan te tonen. Leerlingen zijn dus sterk betrokken bij en op het evaluatieproces. Ook andere kenmerken van alternatieve evaluatievormen kunnen echter bij het gebruik van een portfolio gerealiseerd worden. Zo maken zowel de voortdurende dialoog tussen de leerkracht en de leerling als de leerkansen die een portfolio biedt, de integratie van instructie en assessment mogelijk. Daarnaast kan bij het creëren van een portfolio op een gepaste manier rekening worden gehouden met individuele leerlingenverschillen. De relatie tussen portfolio's enerzijds en constructiegerichtheid en levensechtheid anderzijds is enigszins complexer. Vast staat dat de taken waaruit de leerlingen een selectie moeten maken, voorbeelden kunnen zijn van authentieke vaardigheidstaken. Bovendien kan het creëren van een portfolio op zich aangezien worden als een authentieke vaardigheidstaak. Toch gaat het niet volledig op om te stellen dat een portfolio altijd een perfect voorbeeld is van een vaardigheidsevaluatie. Om de beschrijving van de kenmerken van portfolio's als evaluatie-instrument af te ronden, willen we benadrukken dat portfolio-evaluatie op een viertal manieren kan verschillen van vaardigheidsevaluatie: 1. centraal staat de verzameling van werkstukken en niet zozeer één specifieke prestatie van de leerling; 2. de leerling heeft als het ware meerdere kansen om een goed product te maken dat zijn competenties kan aantonen; 3. er kan een variëteit aan taken in de portfolio gerepresenteerd worden (inclusief elementen als het verwerven van declaratieve kennis); en 4. sommige van de opgenomen stukken zijn afkomstig uit stadia van het leerproces dat men van de leerling nog niet verwachtte dat deze de beoogde doelstelling reeds zou kunnen realiseren. Hierbij denken we onder andere aan kladversies, onvolledig werk, commentaar en zelfkritiek op vroegere taken (Gredler, 1999; George, 1995).

Uit de bovenstaande beschrijving van de kenmerken van portfolio's onthouden we dat de evaluatievorm de mogelijkheid in zich draagt om al de kenmerken van alternatieve evaluatievormen in voldoende mate te realiseren.

3.2 Doelen van een portfolio assessment

Met het gebruik van een portfolio als evaluatie-instrument wil men verscheidene doelen realiseren. De meest belangrijke zijn het uitspreken van beoordelingen over leerlingen en het stimuleren van leerlingen tot reflectie. Het eerste doel omvat het portretteren van de groei en/of de capaciteiten van leerlingen met het oog op een geïnformeerde en doordachte evaluatie. Het tweede omvat het ontwikkelen van de individuele en persoonlijke verantwoordelijkheid van de leerling in het bevragen van en het reflecteren over het eigen werk. Zoals later zal blijken bestaat voor veel gebruikers het belangrijkste doel van een portfolio uit het stimuleren van reflectie bij leerlingen. De portfolio is als het ware een katalysator voor reflectie over de eigen groei. Op die manier is een portfolio niet alleen een evaluatie-instrument maar ook een leerinstrument. De evaluatiemethode kan dus eveneens gebruikt worden om het te onderwijzen curriculum te ondersteunen (Kerka, 1995; ASCD, 1996; Tillema, 1998; Gredler, 1999; Janssens et al., 2000; Rolheiser, et al., 2000).



Even stilstaan... 4.4: Wat wil men bereiken met reflectie?

In een portfolio ligt de nadruk sterk op reflectie. Hiermee wil men onder andere volgende leerdoelen verwezenlijken. Lees deze doelen even door en beantwoord daarna volgende twee vragen.

- ✓ De ontwikkeling van metacognitie (de capaciteit van leerlingen om na te denken over hun denken).
- ✓ De ontwikkeling van een realistische vertrouwen in de eigen capaciteiten als lerende.
- ✓ De ontwikkeling van de zelfevaluatiemogelijkheden van leerlingen (het beoordelen van de kwaliteit van eigen werk op basis criteria).
- ✓ De ontwikkeling van kritisch denken, probleemoplossend gedrag.
- ✓ De ontwikkeling van het dragen van verantwoordelijkheid voor het eigen leerproces.

1. Denk u dat het gebruik van portfolio's ook bij uw leerlingen hiertoe kan bijdragen? Waarom (niet)?
2. Ziet u nog bijkomende of andere doelen?

De combinatie van het proces bij het ontwikkelen van een portfolio enerzijds en de portfolio als product anderzijds maakt dat leerkrachten zich met het gebruik van deze evaluatievorm kunnen richten op een variëteit aan onderwijs- en evaluatiedoelen. Naast het verzamelen van evaluatiegegevens en het stimuleren tot reflectie kunnen dus nog andere doelen nagestreefd worden. In figuur 4.61 beschrijven we een aantal situaties waarin er om andere redenen een beroep wordt gedaan op portfolio's (Office of Research, 1993, Rolheiser et al., 2000).

Figuur 4.61: Andere portfoliodoelen

- ✓ Portfolio's kunnen gebruikt worden om ouders te betrekken bij het onderwijsleerproces van hun kind en om de resultaten van individuele leerlingen naar ouders te rapporteren.

Ouders moeten dan wel vertrouwd gemaakt worden met het verloop van portfolio's en overtuigd worden van de meerwaarde tegenover traditionele toetsen. Eenmaal aan deze voorwaarden voldaan is, blijken ouders het gebruik van portfolio's enthousiast te steunen.

- ✓ Het publiek van een portfolio kan ook bestaan uit toekomstige werkgevers. In dat geval ligt de nadruk op specifieke kennis, vaardigheden en attitudes die verwacht worden in een bepaalde professionele setting.
- ✓ Portfolio's kunnen tevens gebruikt worden om de prestaties van klassen of scholen met mekaar te vergelijken. Wanneer men dergelijke vergelijkingen vooropstelt, moet men speciale aandacht besteden aan het ontwikkelen van criteria en standaarden die de basis zullen vormen van de vergelijking. Ook de vraag wat al dan niet in de portfolio zal worden opgenomen, moet een duidelijk antwoord krijgen.

4 Waarom portfolio's?

In de evaluatieliteratuur worden zowel argumenten beschreven die pleiten voor het implementeren van portfolio's als argumenten die noodzaken tot enige terughoudendheid. In de volgende paragrafen gaan we op beide aspecten kort in.

Leerlingen bewaren al lange tijd hun taken en opdrachten op een systematische wijze. Men kan zich dus de vraag stellen welke vernieuwing een portfolio dan inhoudt. Wat is de meerwaarde ervan als evaluatie-instrument? Essentieel is dat een portfolio leerlingen ertoe aanzet verder na te denken over die systematische verzameling. Portfolio's sporen de leerlingen aan om te reflecteren op hun eigen werk, activiteiten, keuzen en doelen. Precies door leerlingen bij en op het leerproces te betrekken, krijgen ze een beter beeld van zichzelf als lerende. Wanneer leerlingen reflecteren op wat en op hoe ze leren, ontwikkelen ze de vaardigheden die noodzakelijk zijn om zich te ontplooien tot effectieve lerenden. Naast het belang dat wordt gehecht aan reflectie kunnen evenwel nog verscheidene andere redenen gerapporteerd worden die het gebruik van portfolio's bepleiten (Dochy & Segers, 1999; Tondeur, 1998; Office of Research, 1993).

- ✓ *Betrokkenheid en activiteit.* Portfolio's betrekken de leerlingen actief bij de evaluatie. De leerlingen nemen mee verantwoordelijkheid op voor hun eigen leren en vormen een deel van het evaluatieproces. Portfolio's bieden leerlingen een structuur om hun eigen criteria voor goed werk te ontwikkelen en te hanteren. De leerlingen participeren in de selectie van de inhoud en in het samenstellen van de beoordelingscriteria. Deze betrokkenheid en activiteit zijn overigens een noodzakelijke voorwaarde om tot reflectie te komen. Leerlingen kunnen hierdoor hun eigen leerwinsten, hun eigen problemen ten aanzien van bepaalde leerinhouden aanvoelen. Te veronderstellen is dat dit bijdraagt tot de intrinsieke motivatie van de betreffende leerlingen.
- ✓ *Al evaluerend leren.* De leervaardigheden en -strategieën die leerlingen ontwikkelen voor het gebruik van een portfolio zijn wezenlijke vaardigheden om te komen tot levenslang leren. Dit betekent dat leerlingen al evaluerend nieuwe onderwijsdoelstellingen

gen kunnen bereiken. De portfolio is dus een mooi voorbeeld van de integratie van instructie en evaluatie.

- ✓ *Kennis van de verwachtingen.* Het expliciteren van de rubrieken en de criteria maakt dat leerlingen een duidelijk zicht hebben op de doelen en de verwachtingen waaraan ze dienen te voldoen. Leerlingen werken beter als ze deze doelen en verwachtingen kennen. Bij portfolio's kunnen leerlingen bovendien meedenken over wat als leerdoel vooropgesteld wordt.

Bij het invoeren van de portfolio als evaluatie-instrument in Connecticut (VSA) werden een aantal nastrevenswaardige doelen vooropgesteld (ASCD, 1996). De opsomming hiervan is een mooie afsluiter van het pleidooi voor het gebruik van portfolio's. Vermeld werden het verbeteren van de leerlingprestaties; het promoten van vaardigheden zoals zelfevaluatie en doelen vooropstellen; een duidelijk beeld van de leerling als lerende mogelijk maken; een manier introduceren om de vooruitgang van leerlingen te communiceren naar ouders, toekomstige leerkrachten en derden; de verantwoordelijkheid van leerlingen voor hun eigen leren ontwikkelen; ervaringen van controle en fierheid bij leerlingen over de eigen verwezenlijkingen opbouwen en vertrouwen scheppen bij leerlingen ten aanzien van hun mogelijkheden als lerenden. Verschillende redenen dus om het gebruik van portfolio's te bepleiten.

Toch is het gebruik van portfolio's als evaluatievorm is niet steeds vanzelfsprekend en ongenueanceerd positief. Naast de gerapporteerde voordelen zijn er immers een aantal hinderpalen bij het implementeren van portfolio's. Te denken valt aan noodzakelijke voorwaarden zoals de voorkennis van leerlingen. Leerlingen kunnen hun sterktes en zwaktes pas aanvoelen, mits ze over de nodige metacognitieve vaardigheden beschikken (Dochy & Segers, 1999). Het gebruik van portfolio's is bovendien tijds- en arbeidsintensief. Dergelijke elementen zijn echter eerder kenmerkend voor alternatieve evaluatievormen in het algemeen dan voor portfolio's in het bijzonder. Vandaar dat we ze aan bod brengen in deel 5 waar we het invoeren van alternatieve evaluatievormen bespreken.

5 Het evaluatieproces bij portfolio's

In het ontstaan van een portfolio kunnen we drie opeenvolgende fasen onderscheiden, namelijk de organisatie en planning, de eigenlijke verzameling en de reflectie. Elk van deze fasen verdient bijkomende toelichting.

5.1 Organisatie en planning

Net zoals bij andere evaluatievormen is het vooropstellen van de doelen die men met het gebruik van een portfolio wil bereiken, een noodzakelijke stap alvorens men de portfolio effectief vorm wil gaan geven. Volgende vragen kunnen bij het vooropstellen van portfoliodoelen voor enige sturing zorgen: *Waarom implementeer ik een portfolio (opgelegd of vrije keuze)? Waarom*

wil ik leerlingen hun eigen werk laten verzamelen en erover laten reflecteren? Wil ik mij voornamelijk richten op het ontwikkelingsproces, het product als dusdanig of beide? Leg ik de nadruk op de algemene prestatie van de leerlingen of op specifieke domeinen en vaardigheden?

Duidelijkheid over de onderwijs- en evaluatiedoelen zijn echter niet de enige voorwaarden om van start te gaan. De organisatie en planningsfase brengt nog een aantal andere beslissingen vanwege de leerkrachten en/of de leerlingen met zich mee. Door een aantal essentiële vragen uit te klaren voor het begin van het eigenlijke evaluatieproces kan men verzekeren dat de leerlingen duidelijk de doelen en de plaats van de portfolio in het ganse begeleidings- en evaluatieproces kennen. Centrale elementen die op dit punt in het evaluatieproces moeten uitgeklaard worden zijn (Rolheiser et al, 2000):

- ✓ Onderwijsdoelen. Inventariseer alle documenten die een overzicht kunnen geven van de onderwijsdoelen (eindtermen, schooleigen doelen, persoonlijke doelen).
- ✓ Portfoliodoelen. Gezien de onderwijsdoelen, welke zijn de essentiële doelen die je met het gebruik van de portfolio beoogt?
- ✓ Portfoliotype. Welk portfoliotype is het meest aangewezen (beste werk-portfolio of groeiportfolio)? Wat is de focus van de portfolio (vak, curriculuminhoud, vaardigheden)?
- ✓ Participanten. Welke betrokkenen participeren naast de leerkracht en de leerling in het evaluatieproces en welke zijn hun functies (beoordelen, geven van feedback, publiek)?
- ✓ Tijds kader. Over welke periode wordt de portfolio samengesteld?

Een bijkomend element dat we tot de organisatie en planning van de portfolio rekenen, is het bepalen van de manier waarop de portfolio zal worden bijgehouden. De keuze van een klassificeringsmethode is voornamelijk een praktische aangelegenheid. Elk van de boven behandelde elementen moeten hierbij in overweging worden genomen (zoals doelen, portfoliotype, tijds kader en publiek). De keuze van een werkwijze is verder afhankelijk van een aantal praktische feiten en criteria. Volgende twee gerelateerde vragen moeten beantwoord worden om een gepaste klassificeringsmethode te selecteren (Rolheiser et al., 2000).

- ✓ Wat zijn de feiten? Is nog veel ruimte beschikbaar in de klas? Moet de classificatie steeds onmiddellijk beschikbaar zijn? Kunnen leerlingen beschikken over een PC (op school en thuis)? Welke financiële middelen hebben we ter beschikking? Moeten de leerlingen de portfolio transporteren van de school naar huis? Moet de classificatie verschillende mediatypes kunnen omsluiten?
- ✓ Aan welke criteria moet de classificatiemethode voldoen? Deze resulteren onder andere uit bovenstaande feiten. Mogelijke criteria zijn: stevig en robuust, goedkoop, mogelijkheid tot organisatie en tot het bevatten van verschillende soorten werkstukken, makkelijk op te slaan, transporteerbaar, etc.

Elke portfoliocontext zal leiden tot specifieke criteria voor het classificatiesysteem. Op basis van die criteria kan een keuze gemaakt worden tussen onder andere volgende classificatiemethoden.

- | | |
|---------------------|------------------------------|
| (1) een ringmap | (5) gebruikte waspoederdozen |
| (2) kaften | (6) floppy disks |
| (3) een aktetas | (7) een laptop |
| (4) een archiefkast | (8) ... |

Het is aan te bevelen alle organisatorische en planmatige beslissingen systematisch neer te schrijven en te verantwoorden. Op die manier kan later aan de leerlingen en aan andere betrokkenen duidelijk worden gemaakt wat men precies voor ogen heeft. Leerlingen kunnen de structuur wellicht goed gebruiken en bovendien worden er discussies vermeden.

5.2 Verzameling – Samenstellen van de portfolio

De kern van een portfolio-evaluatie is de verzameling van betekenisvolle ervaringen en producten die het leerproces van de leerlingen representeren. Deze verzameling vormt een volgende stap in het portfolioproses. Eerst moeten er echter nog een aantal zaken uitgeklaard worden. Leerlingen moeten bijvoorbeeld kunnen beschikken over richtlijnen die hen duidelijk maken op welke tijdstippen ze best bepaalde werkstukken en materialen selecteren. Ook vragen omtrent hoe ze best de verzamelde items organiseren en presenteren, en hoe de portfolio zal beoordeeld worden, zijn aan de orde alvorens leerlingen aan de slag kunnen.

5.2.1. De inhoud van de portfolio bepalen

In de organisatie van een portfolio is een centrale plaats weggelegd voor de zogenaamde categorieën. In de categorieën wordt uitgedrukt wat belangrijk wordt geacht en waarover de portfolio moet informeren. Het bepalen van categorieën verzekert een gerichte selectie en verzameling van werkstukken (Rolheiser et al, 2000). Categorieën zijn dus nauw verbonden met de onderwijsdoelen. Ze kunnen algemeen of specifiek zijn. Algemene categorieën binnen het vak Nederlands zijn bijvoorbeeld lezen, schrijven en spreken. Specifieke categorieën zijn gedetailleerder en richten zich binnen schrijven bijvoorbeeld op vertellend schrijven, poëzie en zakelijk schrijven of voltooid deelwoorden. Het is belangrijk dat leerlingen beschikken over een duidelijk overzicht van de verschillende categorieën die ze in hun portfolio moeten opnemen. Het is dan ook aan te bevelen de categorieën neer te schrijven in een inhoudstafel of een reflectieformulier.

Figuur 4.62: Voorbeelden van categorieën in een wiskundeportfolio

In een wiskundeportfolio bij de leerlingen van de eerste graad worden bijvoorbeeld volgende categorieën vooropgesteld:

- ✓ Meetkunde
- ✓ Inzicht hebben in de volgorde van bewerkingen
- ✓ Vergelijkingen oplossen met één onbekende
- ✓ Een probleem in een schema weergeven
- ✓ Weergeven van een oplossingsproces

- ✓ Metacognitie:
- Opgaven herformuleren
- Wiskundige problemen schematisch weergeven
- Aandacht hebben voor het oplossingsproces

In dit voorbeeld blijkt dat de specificiteit van categorieën sterk kan variëren. Elk van de categorieën kan bovendien verder gespecificeerd worden. Of dit wenselijk is hangt af van de leerlingenkenmerken en de inhoud van de portfolio.

De categorieën worden door de leerlingen met concrete werkstukken ingevuld. Welke types van informatie hiertoe gebruikt worden sterk kan variëren. De elementen kunnen in een verscheidenheid van vormen worden opgenomen: geschreven, visueel, auditief of multimediaal. Niet enkel wat het gebruikte medium betreft, kan de inhoud echter verschillen, ook de elementen die in de portfolio worden opgenomen, kunnen van zeer uiteenlopende aard zijn. Te denken valt aan volledige werkstukken, kladversies, feedback van de leerkracht of peers, reacties van externen. Kortom: alle informatie die relevant is en voldoende belangrijk geacht wordt om aan te tonen welke vooruitgang een leerling in een vooropgestelde categorie boekte, kan worden opgenomen. Dit geldt bijvoorbeeld ook voor elementen die van belang zijn om het tot stand komen van de portfolio te kunnen begrijpen (zoals de participatie van de leerling in de selectie van de inhoud van de portfolio, de richtlijnen voor deze selectie, criteria ter beoordeling van de verdiensten en verslagen van de zelfreflecties van de student (Dochy en Segers, 1998; Rolheiser, 2000).

Om de inhoud van een portfolio te illustreren presenteren we in figuur 4.63 een aantal voorbeelden van mogelijke werkstukken die in een portfolio kunnen worden opgenomen. Omdat de inhoud van de portfolio verschilt naargelang het type maken we een onderscheid tussen voorbeelden voor een groeiportfolio en een 'beste werk'-portfolio (zie paragraaf 6).

Figuur 4.63: Voorbeelden van de inhoud van portfolio's

Nederlands

- ✓ Bestewerk-portfolio
- Een reactie op een stuk literatuur
- Een origineel gedicht, verhaal, kortverhaal of mythe
- Een script voor een toneelscène, een karakterstudie van een personage
- ✓ Groeiportfolio
- Persoonlijke doelen voor schrijven (gerelateerd aan een sterke- en zwakte analyse)
- Verschillende componenten van een verhaal of essay (plan, kladversies, eindversie)
- Lijst van gelezen boeken

Wiskunde

- ✓ Bestewerk-portfolio
- Oplossing van een prestatietaak
- Een model van een driedimensionale figuur
- Een zelfgemaakte grafiek
- ✓ Groeiportfolio
- Voorbeelden van beste en slechtste prestaties bij een bepaalde taak met bijbehorende reflecties

Evidentie voor het verwerven van verschillende oplossingsstrategieën
Geïllustreerde ontwikkeling van het opstellen van grafieken.

Vakoverschrijdend

- ✓ Bestewerk-portfolio
- Een opname van een presentatie of speech
- Een autobiografische sketch
- ✓ Groeiportfolio
- Voorbeelden van schriftwerk uit verschillende vakken.
- Zelfevaluatie van de vooruitgang bij vakoverschrijdende vaardigheden

5.2.2. Criteria bepalen

Met het bepalen van categorieën weten leerlingen over welke onderwerpen ze gegevens moeten verzamelen. Categorieën geven echter niet aan of een werkstuk dat binnen een categorie past al dan niet voldoende relevant en geschikt is om in de portfolio te worden opgenomen. Het is immers belangrijk dat elk van de in de portfolio opgenomen elementen voldoet aan een aantal criteria en standaarden. Om tot een goede beoordeling van portfolio's te komen moeten dus, zoals bij elke andere evaluatie, criteria ontwikkeld worden. Deze criteria hebben een dubbel doel. Ze ondersteunen enerzijds de selectie van werkstukken die best in de portfolio zullen worden opgenomen en vormen anderzijds de basis voor het beoordelen van de portfolio als geheel. Het ontwikkelen van criteria en rubrieken kan volgens de werkwijze die we voorstelden bij het beschrijven van het evaluatieproces (zie deel 3). Een aantal mogelijke criteria zijn weergegeven in figuur 4.64.

Figuur 4.64: Portfolio-criteria

- ✓ Aantonen van begrip, inzicht en capaciteit;
- ✓ Aantonen van groei en ontwikkeling ten aanzien van de te bereiken onderwijsdoelen;
- ✓ Bedachtzaamheid (ten aanzien van aanwijzingen voor de eigen ontwikkeling en metacognitieve reflectie);
- ✓ Creativiteit en originaliteit;
- ✓ Volledigheid en correctheid;
- ✓ Organisatie van de portfolio;
- ✓ Verscheidenheid van de opgenomen stukken (in die zin dat een bepaalde prestatie of ontwikkeling op verschillende manieren wordt gedocumenteerd);
- ✓ Kritische zelfreflectie.

Leerlingen moeten in ieder geval op de hoogte zijn van welke criteria het onderscheid maken tussen al dan niet relevante informatie. Het beslissen over de criteria (wat gewaardeerd wordt in de leerlingprestaties) en over de standaarden (wat aangezien wordt als een voorbeeldige prestatie voor een bepaald criterium in de specifieke doelgroep) zijn echter op zich leerrijke onderwijservaringen (Birenbaum, 1995). Er wordt dan ook aangeraden de criteria en standaarden gezamenlijk door de leerkracht en de leerlingen te laten ontwikkelen. Tenminste indien de mogelijkheden van de leerlingen dit toelaten (Office of Research, 1993). Dit geldt zowel voor de

criteria die gehanteerd worden ter beoordeling van elke inbreng in de portfolio op zich als voor de portfolio in zijn geheel.

5.2.3. De eigenlijke selectie van werkstukken

Gegeven de verschillende doelen, categorieën en criteria van portfolio's is de verscheidenheid aan mogelijke werkstukken zeer groot. Bij de aard van deze werkstukken en de inhoud van de portfolio staan we stil in paragraaf 7. Hieronder gaan we in op het selectieproces tijdens het creëren van een portfolio. Alvorens men met de eigenlijke samenstelling van de portfolio kan beginnen, moeten immers nog een aantal vragen beantwoord worden. Deze zijn: wie beslist over de inhoud van de portfolio, wanneer wordt er geselecteerd en hoe worden werkstukken geselecteerd?

1. Wie beslist over de inhoud van de portfolio?

De beslissingsvrijheid voor de leerling omtrent wat al dan niet in de portfolio kan worden opgenomen kan sterk variëren. Dit betekent van het vrijlaten van de keuze van wat de leerling zelf relevant acht tot het kiezen op basis van vooraf gespecificeerde categorieën. Essentieel is dat de leerling de verantwoordelijkheid draagt over het selectieproces. Of de uiteindelijke beslissing over het opnemen van een bepaald werkstuk nu al dan niet volledig bij de leerling ligt of door de leerling met de leerkracht (peers of ouders) besproken wordt, is van ondergeschikt belang. Goed om weten is wel dat het overleggen met anderen kan leiden tot interessante feedback en maakt dat leerlingen beter worden in zelfevaluatie (Rolheiser, 2000). Voor de aanvang van de evaluatie moet goed omschreven worden welke de verschillende betrokkenen zijn in het selectieproces en wat de rol van elk van hen is.

2. Wanneer zal de selectie gebeuren?

Ook de vraag wanneer leerlingen de selectie moeten maken, moet vooraf beantwoord worden. Afhankelijk van de doelen van de evaluatie en van de mate waarin leerlingen vertrouwd zijn met het selecteren van werkstukken en het verantwoorden van die keuzes, zal de fasering van het selectieproces variëren. Een aantal mogelijkheden zijn: tijdens vaste momenten tijdens de week of maand, aan het einde van een leerstofonderdeel, aan het einde van een trimester of schooljaar, etc.

3. Hoe selecteren?

Het ontwikkelen van methodes die de leerlingen ondersteunen in het selecteren van werkstukken is een belangrijke taak voor de leerkracht. Het selectieproces kan immers op verschillende manieren verlopen, zowel wat betreft de strategie als wat de tijdsplanning betreft. Ook al zijn de criteria die leerlingen ter beschikking hebben om werkstukken te selecteren op zich voldoende duidelijk, toch is het aangewezen de leerlingen te leren hoe ze de criteria kunnen gebruiken. De

volgende figuur presenteert een aantal richtlijnen om leerlingen op een effectieve manier in te wijden in het selecteren van werkstukken (Rolheiser et al., 2000).

Figuur 4.65: Hoe succesvol leren selecteren?

- ✓ Geef leerlingen zeker de kans om te experimenteren en vertrouwd te worden met het samenstellen van portfolio's. Zorg er voor dat de eerste portfolio die ze samenstellen niet zal gebruikt worden om gewichtige beslissingen op te baseren.
- ✓ Brainstorm met je leerlingen over wat goede voorbeelden zijn om in de portfolio op te nemen. De klassikale voorbeelden leren leerlingen hoe ze individueel te werk moeten gaan.
- ✓ Creëer bij de aanvang van de portfolio voldoende tijd voor leerlingen om de criteria te leren toepassen. Voorzie hen hierbij van voldoende feedback over waarom een werkstuk al dan niet geschikt is en over de reflectie van leerlingen. Niet alleen de leerkracht kan hier een rol spelen, ook medeleerlingen kunnen mekaars selectie van werkstukken beoordelen.

Eenmaal leerlingen vertrouwd zijn met de vooropgestelde criteria en met het gebruik ervan, zal de selectie van werkstukken steeds doelmatiger worden en zal de reflectie op de werkstukken steeds informatiever en leerrijker zijn.

5.3 Reflectie

Reflectie wordt vaak aangezien als het belangrijkste onderdeel van het portfolioproces. We leren immers niet van onze ervaringen, maar van onze reflectie op die ervaringen (Rolheiser et al., 2000). Zonder reflectie zou een portfolio enkel een doelloze verzameling zijn van het werk van leerlingen. Waar mogelijk moeten in de portfolio dan ook gegevens worden opgenomen over de reflecties van leerlingen. Hierbij richt men zich zowel op het verwerven van kennis en (metacognitieve) vaardigheden als op persoonlijke ervaringen en bemerkingen. Zoals in het hoofdstuk 'betrokkenheid van leerlingen' bleek, kunnen dergelijke reflecties verscheidene vormen aannemen (zoals logboeken, vragenlijsten). Centraal staan altijd vragen in de zin van:

- ✓ Wat heb ik gedaan? Wat heb ik geleerd?
- ✓ Welke gevoelens heb ik hierbij?
- ✓ Waarom heb ik dit document in mijn portfolio opgenomen?
- ✓ Wat zou ik volgende keer beter willen doen?

Naast de verschillende werkstukken die in de portfolio worden opgenomen is dus ook de reflectie een essentieel onderdeel. Deze reflectie kan zich zowel richten op de portfolio in zijn geheel als op een specifiek werkstuk in het bijzonder. Beide reflecties worden kort geïllustreerd. Voor een beschrijving van de verschillende manieren waarop reflectie binnen portfolio's kan worden gerealiseerd, verwijzen we naar de zelfevaluatieinstrumenten die we in hoofdstuk 3 van dit deel presenteerden.

Figuur 4.66: Reflectie over gans de portfolio

Korte introductie bij een taalportfolio

Bij de aanvang van de lessen werd mij gevraagd om stil te staan bij mijn sterke en zwakke punten en hoe ik de lessen Nederlands tegemoet zag. Bij het nalezen van deze evaluatie aan het einde van het schooljaar, merkte ik dat ik op alle domeinen vooruitgang heb geboekt. Voornamelijk het feit dat ik ben afgestapt van mijn neiging om schrijftaken uit te stellen maakt me zeer blij. Ik ben erin in geslaagd alles tijdig af te werken. Dat gaf me meer tijd om nog verbeteringen aan te brengen. Ondanks het feit dat ik vooruitgang heb geboekt, kan ik echter nog blijven groeien als schrijver en kan ik nog harder werken om goede prestaties neer te zetten. Wat ik met deze laatste twee bemerkingen bedoel, zal ik duidelijk maken doorheen mijn portfolio.

Figuur 4.67: Reflectie bij een specifiek werkstuk

Volgende vragen worden gebruikt als leidende vragen bij leerlingen die een zekere structuur nodig hebben bij het reflecteren over specifieke werkstukken. Afhankelijk van de zelfevaluatie-capaciteiten van de leerlingen moeten de sturing en de moeilijkheid van de vragen aangepast worden.

- ✓ Ik ben zeer trots op dit werk omdat...
- ✓ Wat ik leerde door aan dit stuk te werken is...
- ✓ Dit werkstuk was uitdagend omdat het...
- ✓ Dit werkstuk toont aan dat ik weet hoe te...
- ✓ Dit werkstuk is belangrijk omdat het...
- ✓ Ik verwacht dat dit werk indruk zal maken op mijn ouders omdat...
- ✓ In welk opzicht ben je gegroeid of veranderd?
- ✓ Waarom heb je dit werkstuk opgenomen?

Een reflectie over een werkstuk dat in de portfolio werd opgenomen kan er bijvoorbeeld als volgt uitzien. Aan het woord is Stef een leerling uit de eerste graad SO die de selectie van een werkstuk in zijn schrijfportfolio verantwoordt.

Dit werkstuk is het resultaat van de opdracht om op basis van een foto een verhaal te schrijven. Ik heb dit stuk gekozen omdat ik er veel tijd aan heb besteed en omdat ik aanvankelijk wat moeite had om van de foto tot een verhaal te komen. Toch vond ik het een zeer leuke en leerrijke oefening om de mensen op de foto tot leven te brengen. Het resultaat mag er volgens mij wezen. Ik heb deze taak aangepakt zoals door de leerkracht werd voorgesteld. Vroeger ging ik niet zo gestructureerd te werk. In tegenstelling tot bij vroegere opdrachten ben ik niet onmiddellijk begonnen met schrijven. Eerst heb ik in een schema de relaties tussen de karakters aangegeven en goed nagedacht over een aantal mogelijke plots. Vroeger koste mijn ongestructureerde aanpak minstens evenveel energie terwijl het resultaat minder was.

6 Verschillende types portfolio

In de literatuur worden verschillende types portfolio's onderscheiden die elk specifieke doelstellingen beogen. Deze verscheidenheid aan portfoliotypes resulteert in verschillende categori-

eringen (Birenbaum, 1995; Dochy & Segers, 1998; Gredler, 1999; Rolheiser, 2000). Het meest interessant is evenwel de eenvoudige onderverdeling in 'beste werk'- en groeiportfolio's.

a. De 'beste werk'-portfolio (exemplarische portfolio, showcase portfolio)

Deze portfolio is een verzameling van voorbeelden die de beste en de meest kwaliteitsvolle werkstukken van de leerling samenbrengt. De showcase portfolio bevat alleen de beste producten die de leerling tijdens het leerproces voortbracht en bevat dus alleen afgewerkte items. De nadruk ligt op wat de leerling kan. Een dergelijke portfolio kan beschouwd worden als een verzameling waarin men toetsen, projecten en verbeterd materiaal opneemt die het meest accuraat de 'beste prestatie' van leerlingen reflecteren. Dat materiaal kan zowel betrekking hebben op het product van het leren als op het leerproces zelf.

b. De groeiportfolio

De groeiportfolio bevat voorbeelden van het ontwikkelingsproces van de leerling en bestaat zowel uit afgewerkte als niet-afgewerkte stukken, uit effectieve en niet-effectieve elementen, uit kladversies, verslagen van brainstorming en uit andere relevante elementen omtrent het leerproces van de leerling. De nadruk ligt op de ontwikkeling die de leerling doormaakt. De groeiportfolio kan hiertoe zowel een korte als een lange (een geheel schooljaar) periode overkoepelen. Zowel indicaties voor succes als voor mislukking worden in de portfolio opgenomen. Zeer belangrijk bij de groeiportfolio is tevens de reflectie van de lerende waarin deze aangeeft wanneer er groei is opgetreden en waaraan deze groei is toe te schrijven.

Beide groepen portfolio's zijn pas betekenisvol wanneer bepaalde items geselecteerd worden om bepaalde educatieve ervaringen, ontwikkelingen en doelen in de verf te zetten. In beide gevallen is de reflectie van de leerling dus essentieel. In het eerste geval is de kern van de reflectie waarom de opgenomen stukken volgens de leerling als beste stukken fungeren, in het tweede geval is de kern waarom de opgenomen stukken representatief zijn voor de groei die de leerling doormaakte.

7 De verschillende componenten van een portfolio

Doorgaans wordt er een onderscheid gemaakt tussen twee componenten in een portfolio. Centraal staan de werkstukken van leerlingen die worden opgenomen met de bijbehorende reflecties. Daarnaast moet echter eveneens aandacht worden besteed aan de structuur van de portfolio. De verduidelijking van de organisatie van de portfolio kan in principe worden losgekoppeld van de eigenlijke inhoud van de portfolio. Vandaar het onderscheid tussen inhoudelijke en organisatorische componenten.

7.1 Inhoudelijke componenten

Om een inhoudelijk kwaliteitsvolle portfolio samen te stellen zijn drie componenten van belang, namelijk diepgang, breedte en reflectie. **Diepgang** heeft betrekking op de nood aan een goed gedocumenteerde weergave van de evolutie en de prestaties van de leerling. In het geval van een schrijfportfolio kan diepgang bijvoorbeeld bestaan uit nota's, schema's, kladversies en een finale versie van een verhaal. Er moeten met andere woorden voldoende gedetailleerde gegevens aangebracht worden om het leren te documenteren. Onder **breedte** verstaan we dat verscheidene indicaties worden opgenomen van wat men wil documenteren. In het geval van een geschiedenisportfolio die zich toelegt op het leven van soldaten tijdens WO II kan deze component bijvoorbeeld bestaan uit gelezen krantenknipsels, fragmenten uit relevante boeken, officiële staatsdocumenten, getuigenissen van oud-strijders, reacties van de leerling op bepaalde situaties, etc. Hoeveel werkstukken moeten worden opgenomen om voldoende diepgang en breedte te realiseren is afhankelijk van de evaluatiedoelen en kenmerken van de leerlingen. Het is best mogelijk dat bijvoorbeeld drie of vijf werkstukken volstaan. De derde belangrijke component bestaat uit **reflecties** op en analyses van het werk van de leerling. Deze component bestaat voornamelijk uit een beschrijving door de leerling zelf van de belangrijkste kenmerken van het eigen werk, van de eigen sterktes en zwaktes, van veranderingen over de tijd en de vooruitgang die in de toekomst nog moet gemaakt worden. Ook reflecties en feedback van de leerkracht of van andere betrokkenen behoren tot deze derde component.

7.2 Organisatorische componenten

Een portfolio kan inhoudelijk op verschillende manieren georganiseerd en gestructureerd worden. Te denken valt aan volgende indelingen: per onderwerp, per doelstelling, chronologisch, per type product of naargelang de leerling de stukken belangrijk acht voor de eigen ontwikkeling. In principe kunnen ook zeer verscheidene redenen aan de basis liggen van de selectie van portfoliostukken. Bij wijze van voorbeeld vermelden we de volgende: ik heb er hard voor gewerkt, ik heb er veel van geleerd, het was interessant, de taak daagde me uit en ik heb het goed voorbereid. Deze achterliggende principes moeten steeds duidelijk in een begeleidend schrijven/deel worden toegelicht.

Om het samenstellen en lezen van de portfolio te structureren, is het daarom belangrijk dat een duidelijke leidraad wordt opgenomen. De lezer van de portfolio moet op een snelle en eenvoudige manier een overzicht kunnen krijgen van de hoeveelheid en de aard van de opgenomen werkstukken, van het curriculumdomein waartoe de stukken behoren, van de frequentie van zelf- en peerevaluaties, van waar reflecties kunnen gevonden worden, etc. Naast structuur is ook het aanbieden van achtergrondinformatie over de leerling en de onderwijscontext van belang. Volgende werkwijzen (of combinatie daarvan) kunnen hiertoe worden aangewend (Rolheiser et al., 2001; Mabry, 1999).

1. Inhoudstafels

Een chronologische inhoudstafel van de werkstukken die de leerling in de portfolio toevoegt, is een eerste mogelijkheid om het overzicht in de portfolio te bewaren. Telkens de leerling iets toevoegt, wordt dit in de inhoudstafel aangegeven (datum, titel, kern van het werkstuk). De inhoudstafel kan dus tijdens de samenstelling van de portfolio veranderen (leerlingen werken dan best met steekkaarten of op een PC). Bij de finale selectie van werkstukken kan de leerling een vaststaande inhoudstafel opstellen. Deze hoeft dan niet chronologisch gestructureerd te zijn. Inhoudstafels zijn vooral geschikt om de structuur van de portfolio te duiden. Volgende werkwijzen zijn dan weer beter geschikt om ook achtergrondinformatie te verschaffen.

2. Visuele leeswijzers

Leerlingen kunnen grafische representaties maken om de essentiële aspecten in hun portfolio te structureren. Hierbij kunnen zowel persoonlijke elementen over de leerling als gegevens over de onderwijscontext worden opgenomen. De mogelijkheden zijn zeer ruim: tekeningen, mentale schema's, stroomdiagrammen, boomstructuren. Een leerling kan bijvoorbeeld zichzelf tekenen op een brug tussen zijn vroegere loopbaan en de nog af te leggen leerweg. Op die tekening kan hij aangeven wat er reeds achter zich ligt (wat hij reeds beheerst, waar hij goed in is) en waar hij naar toe wil (doelstellingen, plannen). Ziet de leerweg er uit als een voortdurende steile klim naar een bergtop die zich nog in de wolken bevindt of eerder als vrolijke wandeling door een glooiend landschap met een aantal onverwachte valkuilen. Ook de vergelijking met een polsstokspringer kan de portfolio structureren. Welke vaardigheden wil de leerling bereiken? Hoe zit het met de aanloop en de hulpmiddelen (de polsstok)? Welke obstakels moeten overwonnen worden (zwaartekracht, angst om niet zacht neer te komen)?

3. Geschreven introducties

Een derde werkwijze om de lezer wegwijs te maken in de structuur van een portfolio is het schrijven van een introductie. De geschreven introductie richt zich naar het publiek van de portfolio en geeft de lezer tevens een eerste impressie van de portfolio. Er moet dus zorgvuldig werk van worden gemaakt. Aan het einde van het portfolioprocess –net voor de portfolio zal worden afgewerkt– moeten de leerlingen daartoe kunnen beschikken over voldoende tijd en ondersteuning. Die ondersteuning kan bijvoorbeeld bestaan uit het vooropstellen van een aantal vragen die in de introductie zouden kunnen beantwoord worden (Rolheiser et al., 2000). Deze zijn bijvoorbeeld:

- ✓ Verduidelijk de inhoud en het doel van de portfolio. Ga ervan uit dat de lezer niet bekend is met portfolio's als evaluatieinstrumenten.
- ✓ Beschrijf hoe de portfolio georganiseerd is en beargumenteer die keuze.
- ✓ Welke vind je zelf de beste werkstukken en waarom?
- ✓ Wat vind je uitdagend aan het samenstellen van een portfolio, en wat niet?
- ✓ Waarom zou je het gebruik van portfolio's aan andere leerlingen aan- of afraden?
- ✓ Heb je de portfoliодоelen volgens jou bereikt?

Ter afsluiting van de beschrijving van de verschillende componenten zetten we in de volgende figuur de inhoud van een portfolio op een rij.

Figuur 4.68: De verschillende componenten van een portfolio

Uit de beschrijving van de inhoudelijk en organisatorische componenten leren we dat een portfolio structureel doorgaans bestaat uit volgende elementen (Janssens et al., 2000):

1. een inhoudstafel, leeswijzer en/of introductie (die aangeeft wat in de portfolio terug te vinden is);
2. een beschrijving van het gepresteerde werk;
3. een beschrijving van de concrete leerervaringen en van de evolutie van het leerproces;
4. een selectie van documenten en werkstukken als bewijsmateriaal;
5. een reflectie van de leerling op de vorige elementen (met eventueel bemerkingen van de leerkracht of derden op de portfolio in het algemeen en op deze reflectie in het bijzonder).

8 Het beoordelen van een portfolio

De portfolio in zijn geheel biedt meer informatie dan de som van de onderlinge delen. Toch zijn zowel het geheel als de verschillende delen van belang bij het beoordelen. De evaluatie vereist een zorgvuldige constructie van beoordelingscriteria en het definiëren van wat men verstaat onder een zwakke, adequate en sterke invulling van elk van de criteria. Om een portfolio te beoordelen, moet men duidelijk expliciteren wat men wil beoordelen. Dit betekent dat voor de evaluatie van de portfolio zowel criteria ter evaluatie van de gehele portfolio als criteria ter evaluatie van de verschillende constituerende delen moeten vooropgesteld worden. Het ontwikkelen van beide sets van criteria is echter niet vanzelfsprekend.

Zoals bij andere vormen van evaluatie kan men op basis van de evaluatiecriteria rubrieken samenstellen. Deze kunnen sterk variëren naargelang de benadrukte criteria en naargelang het uitzicht en de uitgebreidheid. Getuige hiervan volgende twee voorbeelden.

Figuur 4.69: Holistische rubriek voor het beoordelen van een portfolio

- (10) – De portfolio is goed gedocumenteerd en georganiseerd. De structuur is accuraat, compleet en makkelijk te volgen. Een excellente weldoordachte portfolio die aantoont dat er een uitstekende inspanning is geleverd.
- (8) – Redelijk goed gedocumenteerd en georganiseerd. De structuur is meestal accuraat, compleet en makkelijk te volgen. Een goede, kwaliteitsvolle portfolio die getuigt van een excellent inspanning.

(6) – Bepaalde delen zijn zwak of inaccuraat gedocumenteerd en/of georganiseerd. De portfolio is niet volledig en niet eenvoudig te volgen. Een gemiddelde, adequate portfolio die aantoont dat er een goede inspanning is geleverd.

(4) – Ongeorganiseerde, slecht gedocumenteerde portfolio. Grote delen zijn onvolledig en/of inaccuraat. Moeilijk te volgen. Een zwakke kwaliteit die getuigt van een zwakke inspanning.

(2) – De portfolio is onvolledig, foutief en/of inadequaar. Het product getuigt van weinig of geen inspanning.

(0) – De portfolio werd niet ingediend.

Figuur 4.70: Analytische rubriek voor het beoordelen van een portfolio

	4	3	2	1	0
Inhoud	De pf bevat al het vereiste materiaal	De pf bevat het meeste van het vereiste materiaal	De pf bevat een deel van het vereiste materiaal	De pf bevat weinig van het vereiste materiaal	De pf bevat enkel irrelevant materiaal
Organisatie	De pf is volledig en afgewerkt georganiseerd. De lezer heeft geen enkele moeite alles te vinden.	De pf is goed georganiseerd. De lezer heeft weinig moeite alles te vinden.	De pf is eerder goed georganiseerd. De lezer kan wel enkele problemen ondervinden om alles te vinden.	De pf getuigt van een poging tot organisatie. De lezer ervaart moeilijkheden alles te vinden.	Is er geen poging tot organisatie. Het is een verzameling van losse stukken.
Kwaliteit	De pf bevat werk van uitzonderlijke kwaliteit. Het is beter en verdergaand dan vereist.	De pf bevat werk van aanvaardbare kwaliteit en komt tegemoet aan de verwachtingen.	De pf bevat werk van aanvaardbare kwaliteit en komt tegemoet aan de meeste verwachtingen.	De pf bevat werk dat eerder van onaanvaardbare kwaliteit is. Verschillende vereisten zijn onvold.	De pf bevat werk dat niet van een aanvaardbare kwaliteit is.

Deze rubrieken bevatten enkel criteria die betrekking hebben op de kwaliteit van de portfolio zelf. Daarom lijkt het alsof de portfolio zonder specifieke leerdoelen wordt samengesteld. Uiteraard is het in principe mogelijk een portfolio bij te houden waarin leerlingen reflecteren over hun leerproces betreffende het samenstellen van een portfolio. De nadruk op de portfolio zelf in evaluatierubrieken is vaak nog groter wanneer leerlingen verwacht worden een elektronische portfolio samen te stellen. In die gevallen worden in de rubriek als criteria meestal ook elementen als voorstelling, navigatie, gebruik van media en dergelijke opgenomen.

In realiteit worden portfolio's echter verzameld om het leerproces van leerlingen in schrijven, lezen, rekenen of aardrijkskunde te documenteren. Om die prestaties te beoordelen kan men geen beroep doen op bovenstaande rubrieken. Doorgaans zal men bij het beoordelen van portfolio's daarom twee (groepen) rubrieken gebruiken. Eén om de eigenlijke leerprestaties (de verschillende werkstukken) te evalueren en één om de portfolio op zich te beoordelen.

9 Voorbeelden van portfoliogebruik

Het ontwikkelen van voorbeelden ligt anders bij portfolio-evaluatie dan bij de andere vormen van alternatieve evaluatie. Uit de bovenstaande beschrijving bleek dat een succesvol gebruik van portfolio's in belangrijke mate een beroep doet op de reeds aangehaalde voorbeelden van (authentieke) vaardigheidstaken en leerlingenbetrokkenheid. Bij deze voorbeelden van taken, opdrachten en beoordelingsinstrumenten gaan we hier niet opnieuw stilstaan. Ook de criteria die gebruikt kunnen worden om portfolio's te beoordelen werden reeds kort toegelicht en geïllustreerd. Vandaar dat we ons er in dit punt toe beperken om voor de drie doelgroepen kort een aantal mogelijkheden tot portfoliogebruik aan te halen.

Doorheen de beschrijving van portfolio's zijn reeds een aantal voorbeelden van taalportfolio's aangehaald. De taalkundige taken en producten van de leerlingen uit de eerste graad lenen zich vaak uitstekend tot portfoliogebruik. Op welk schrijfproduct zijn ze trots, op het welk minder? Waarom vinden leerlingen het ene verslag kwaliteitsvoller dan het andere? Kunnen ze aan de hand van een selectie van hun taken aantonen dat (en wat) ze gaandeweg bijgeleerd hebben? Kunnen ze met schrijffoorbeelden uit andere vakken dan Nederlands demonstreren dat ze in het algemeen bereid zijn om hun handschrift en lay-out te verzorgen?

Gelijkaardige leer- en evaluatiekansen kunnen gerealiseerd worden door het gebruik van portfolio's in het wiskundeonderwijs. Uit welke taken blijkt dat de leerling gaandeweg meer inzicht heeft gekregen in het verband tussen de verschillende bewerkingen of in het oplossen van vergelijkingen? Welk schema van een wiskundig probleem is werkelijk een modelvoorbeeld? Welke voorbeelden tonen aan dat de probleemoplossingsmethode van de leerling veranderd is? Wat ontbreekt er nog aan? Welke oplossing bewijst dat de leerling de vereiste doelstellingen onder de knie heeft?

Ook bij leerlingen Elektriciteit/elektronica komen verschillende producten en taken in aanmerking als basis voor een portfolio-evaluatie. Welke kringloop die de leerling samenstelde, voldoet het best aan de veiligheidsvoorschriften? Waarin verschilt het beste verslag over uitgevoerde metingen dat de leerling samenstelde van het minst goede? Uit welke taken blijkt dat de leerling op een efficiënte manier meettoestellen kan gebruiken? Gesteld dat de leerling moet solliciteren bij een toekomstige werkgever: welke taak toont aan dat de leerling op een gestructureerde manier fouten in een signalisatie schakeling kan aanpakken.

Portfoliogebruik hoeft zich niet te beperken tot de (zogenaamd) sterkere richtingen. Ook bij leerlingen uit het beroepsonderwijs kan de evaluatievorm succesvol worden angewend. De vakken 'Omgangskunde' en 'Praktijk verzorging' in de richting Verzorging verschillen van andere richtingen in die zin dat de activiteiten vaak in mindere mate resulteren in producten (in vergelijking met briefwisseling in de richting kantoor, een lasverbinding of een zelf gemaakt kledingsstuk). Toch zijn er ook binnen deze studierichting interessante mogelijkheden. Leerlingen kunnen een selectie maken uit hun stageverslagen en verantwoorden waarom het gekozen verslag hun 'beste' prestatie is. Zijn ze geëvolueerd in de activiteiten die ze organiseren voor

bejaarden? Hoe blijkt dit uit hun activiteiten? Welke stage-ervaringen maken volgens de leerling zelf dat hij/zij (nog niet) klaar is om als verzorgende aan de slag te gaan?

Deel



5

Aanvangen met alternatieve evaluatievormen

INHOUD - DEEL 5

Aanvangen met alternatieve evaluatievormen

<i>Inleidend</i>	271
<i>Over de opbouw van dit deel</i>	272
Hoofdstuk 1: Valkuilen op het pad naar alternatieve evaluatievormen	272
1 <i>Tegenzin van de betrokkenen om nieuwe evaluatievormen in te voeren</i>	273
2 <i>Blijvende gerichtheid op de instructie</i>	275
3 <i>Problemen bij het ontwikkelen en gebruiken van evaluatiecriteria</i>	275
4 <i>Tijdsintensief karakter van vernieuwingen</i>	276
5 <i>Onvoldoende ondersteuning en training van leerkrachten</i>	277
6 <i>Problemen bij de leerlingen</i>	278
Hoofdstuk 2: Hoe valkuilen vermijden?	278
1 <i>Beleidsmakers</i>	279
1.1 <i>Realiseren van een mentaliteitswijziging</i>	279
1.2 <i>Opleiden van (aspirant) leerkrachten</i>	280
1.3 <i>Faciliteren van het invoeren van alternatieve evaluatievormen</i>	281
2 <i>Leerkrachten</i>	282
2.1 <i>Begin kleinschalig en voorzie voldoende tijd</i>	282
2.2 <i>School jezelf in het uitvoeren van goede en alternatieve evaluatievormen</i>	283
2.3 <i>Leer uit commentaren, kritieken en eigen fouten</i>	284
2.4 <i>Communiceer</i>	284
2.5 <i>Maak leerlingen vertrouwd met alternatieve evaluatievormen</i>	285
2.6 <i>Voorzie onderlinge ondersteuning en betrokkenheid in het ontwikkelingsproces</i>	286
2.7 <i>Herverdeel de beschikbare tijd</i>	286
3 <i>Leerlingen</i>	287
4 <i>Ouders</i>	288
Samenvatting	289

Inleidend

Vele leerkrachten die voorheen niet vertrouwd waren met de evaluatievormen die in dit boek beschreven werden, zullen het gevoel hebben dat het onmogelijk is een dergelijke assessmentcultuur met de voorgestelde alternatieve vormen in te voeren in hun school, in hun klas of in hun vakgebied. Aanvangen met alternatieve evaluatievormen is -toegegeven- niet vanzelfsprekend. Het is in zekere mate tijdrovend en dus veel kostend. Leerkrachten, beleidsmensen, ouders en leerlingen hebben de tijd nodig om de kanteling naar een assessmentcultuur te maken en om een betekenisvol en authentiek evaluatiesysteem te creëren. Toch zijn deze kosten enigszins misleidend (Wiggins, 1990). De voordelen en positieve gevolgen van een evaluatiepraktijk die zoveel mogelijk tracht tegemoet te komen aan de eisen van alternatieve evaluatievormen zijn immers talrijk. Dit neemt niet weg dat over het aanvangen met (of het implementeren van) alternatieve evaluatievormen duidelijk en correct moet gerapporteerd worden.

Welke zijn mogelijke obstakels die moeten overwonnen worden? Welke factoren maken dat een implementatieproces al dan niet succesvol zal verlopen? Hoe kan men de implementatie faciliteren? Deze vragen staan centraal in dit laatste deel.

Eén van de belangrijkste uitdagingen die een kanteling naar een assessmentcultuur met zich meebrengt, is een mentaliteitsverandering die de traditionele visie op evaluatie en op de rol van leerkrachten en leerlingen hierin achter zich laat. Maar -zoals reeds herhaaldelijk benadrukt- moet niet alleen evaluatie maar ook de relatie van evaluatie met de instructie en het leren van leerlingen herzien worden. Dit betekent dat bij het invoeren van alternatieve evaluatievormen niet alleen de evaluatiefase onder de loep moet worden genomen. De overgang naar alternatieve evaluatievormen omhelst de noodzaak van het herdefiniëren van de onderwijsdoelen en de instructiemethodes. Om succesvol gebruik te maken van de voorgestelde evaluatievormen moeten -waar nodig- ook de onderwijsdoelen en de instructie in het implementatieproces van de evaluatie worden opgenomen. Het is belangrijk om al de veranderde opvattingen te onderzoeken en te bediscussieren.

Wanneer opvattingen over leren en onderwijsdoelen evolueren moeten evaluatiemethodes worden aangepast aan de hieruit voortvloeiende veranderingen (Fisher & King, 1995). Herinner u wat we in deel 1 schreven over de samenhang tussen leren, onderwijzen en evalueren. Een duidelijke en door de gehele schoolgemeenschap gedragen visie op de doelen en de rol van het onderwijs is dus noodzakelijk om een gezamenlijk evaluatiebeleid uit te bouwen. Zonder overeenstemming over welke kansen de school aan leerlingen moet bieden, en over de rol van de leerkracht in het leerproces is het zeer moeilijk om bestaande opvattingen en onderwijsdoelen te gebruiken als leidraad voor het opzetten van gepaste evaluatiemethodes. Pas in dat geval kan de evaluatie immers een instrument worden dat op haar beurt het klasgebeuren richting geeft, en pas dan kunnen leerkrachten een deel van hun instructietijd 'opofferen' om meer tijdrovende evaluatievormen aan te wenden. Aanvangen met een onderwijscultuur die op een gepaste manier gebruik maakt van alternatieve evaluatiemethoden vergt een mentale verandering en dus

tijd (Fisher & King, 1995). Er is sprake van een cultuurwijziging in die zin dat zowel het denken over onderwijs en evaluatie als de relatie tussen evalueren en instrueren fundamenteel moeten veranderen. Zowel leerkrachten, leerlingen, ouders als beleidsvoerders hebben daarom nood aan een voldoende lange tijdsperiode om de noodzakelijke veranderingen door te voeren. In dit deel stellen we een aantal praktische suggesties voor die een stapsgewijze realisatie hiervan mogelijk maken.

Over de opbouw van dit deel

Ondanks het belang van de onderwijsactiviteiten die aan de eigenlijke evaluatie voorafgaan, zullen we ons in het verdere verloop van dit deel richten op het implementeren van alternatieve evaluatievormen. Dit impliceert niet dat er geen aandacht meer zal zijn voor andere onderwijsactiviteiten (zoals planning en instructie) maar wel dat deze steeds vanuit het perspectief van de evaluatie zullen benaderd worden.

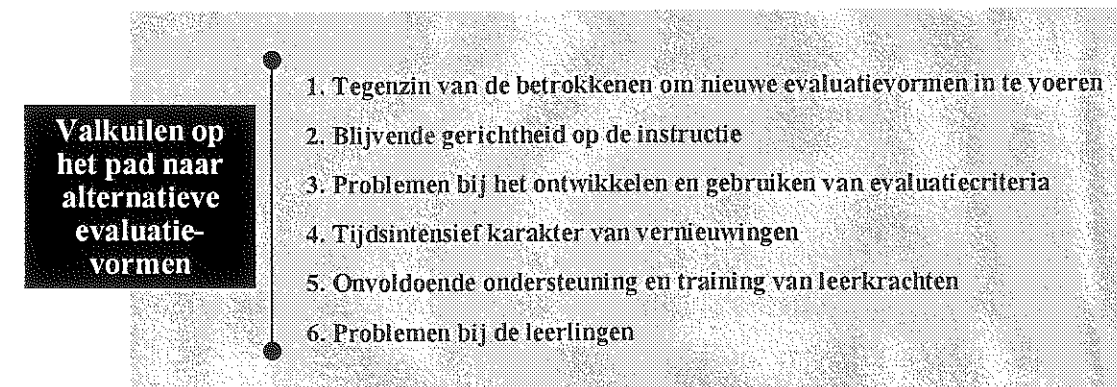
In het volgende staan we eerst stil bij een aantal problemen die kunnen opduiken bij het implementeren van alternatieve evaluatievormen. We gaan op zoek naar welke valkuilen zich kunnen voordoen op het pad naar alternatieve evaluatievormen. Daarna richten we ons op wat de verschillende betrokkenen in een evaluatiecontext kunnen doen om toch succesvol te kunnen aanvangen met een vorm van alternatieve evaluatie. Of met andere woorden: hoe kunnen we de valkuilen vermijden of zoveel mogelijk teniet doen. We zullen hierbij een onderscheid maken tussen de rol van de leerkracht, de rol van beleidsmakers, de rol van de leerling en de rol van ouders. Het implementeren van alternatieve evaluatievormen is immers een gezamenlijk proces waarin de verschillende betrokkenen elk een taak te volbrengen hebben.

Hoofdstuk 1: Valkuilen op het pad naar alternatieve evaluatievormen

De overgang van een traditionele evaluatiecultuur naar een assessmentcultuur is geen magische gebeurtenis die van de ene dag op de andere kan gerealiseerd worden. Het hoeft daarentegen ook geen moeilijk en aanslepend proces te zijn. Althans niet indien we ons bewust zijn van een aantal mogelijke valkuilen en indien we leren uit reeds bestaande ervaringen. Net als elke andere implementatie van een onderwijskundige vernieuwing vereist het gebruik van alternatieve evaluatievormen een herstructurering van het onderwijsleerproces en van de manier waarop leerlingen en leerkrachten interageren. Het implementeren van alternatieve evaluatievormen vereist met andere woorden dat een aantal hinderpalen overwonnen worden (Wolfe & Miller, 1997; Dochy & Segers, 1999). Het bestaan van dergelijke hinderpalen maakt dat leerkrachten, leerlingen en beleidsmensen ontmoedigd kunnen worden en dat de kans op het slagen van de

vernieuwing kleiner wordt. In de volgende paragrafen zetten we een aantal potentiële problemen naast elkaar. Deze zijn samengebracht in de volgende figuur.

Figuur 5.1: Potentiële valkuilen



Gezien de recentheid van de ontwikkelingen aangaande het gebruik van alternatieve evaluatievormen is onderzoek naar mogelijke obstakels nog eerder beperkt (Wolf & Miller, 1997). Toch kunnen we belangrijke lessen trekken uit een aantal ervaringen van collega's die het implementatieproces reeds doorlopen hebben (Westerberg, 1996; Fullan, 1996). Daarnaast kunnen we putten uit de talrijke ervaringen met veranderingsprocessen in het algemeen. Het invoeringsproces van andere onderwijskundige vernieuwingen is immers reeds door talrijke praktijkmensen doorlopen en door verscheidene onderzoekers bestudeerd. Toch gaan we hier niet in op algemene theorieën of strategieën voor verandering. In het volgende destilleren we uit de beschikbare informatie relevante aandachtspunten en kritische factoren om succesvol met alternatieve evaluatievormen aan te vangen. We richten ons met andere woorden specifiek op die aspecten die bij het invoeren van alternatieve evaluatievormen van belang zijn (concreter: hoewel bijvoorbeeld ook leiderschap bij de implementatie van alternatieve evaluatievormen van belang is, zullen we het hier niet aan bod brengen).

1 Tegenzin van de betrokkenen om nieuwe evaluatievormen in te voeren

Niet alle betrokkenen in het onderwijs- en evaluatiegebeuren zijn als vanzelfsprekend enthousiast over de stap naar alternatieve evaluatievormen. Problemen met collega's of andere betrokkenen die het gebruik van alternatieve evaluatievormen niet begrijpen en ondersteunen zijn niet uit de lucht gegrepen. De traditionele onderwijs- en evaluatieculturen zijn sterk ingeburgerd zodat alternatieve evaluatievormen door sommigen ervaren worden als een tijdelijke trend waarvan de noodzaak niet onderschreven wordt. Hierin schuilt wellicht de belangrijkste valkuil op het pad naar alternatieve evaluatievormen. Een mentaliteit- en cultuurverandering zijn dan ook onontbeerlijk om succesvol (d.i. vlot en blijvend) de stap naar het gebruik van alternatieve evaluatievormen te kunnen zetten. De mentale kanteling vormt dan ook de belangrijkste uitdaging voor zowel leerkrachten, ouders, leerlingen als beleidsmakers en andere betrokkenen. De tegenzin van betrokkenen is vaak een gevolg van een gebrek aan eensgezindheid of onduidelijkheid over de doelen die men met de invoering van een vernieuwing wil realiseren. Het is dan ook van groot belang dat men de invoering van alternatieve evaluatievormen begint met een duidelijk doel voor ogen. Hiertoe moet één centrale vraag beantwoord worden: *wat willen we dat onze leerlingen kennen en kunnen wanneer ze onze school verlaten.*

In volgende denkoefening staan we even stil bij de vraag of er in uw school sprake is van een gemeenschappelijke visie op leren, onderwijzen en evalueren.

Even stilstaan... 5.1: De samenhang tussen onderwijsdoelen, instructie en evaluatie

- ✓ Stelt u zichzelf de bovenstaande vraag af en toe?
- ✓ Bespreekt u het antwoord op deze vraag wel eens met uw collega's?
- ✓ Is er overeenstemming bij de leerkrachten over wat goed onderwijs is en over de doelen die nagestreefd worden? Zitten uw collega's (en uzelf) op dezelfde golflengte omtrent wat het eindresultaat zou moeten zijn van het onderwijs?
- ✓ Wordt er in uw school wel eens gereflecteerd over de vraag of de gangbare werkvormen effectief zijn in het bereiken van de vooropgestelde doelstellingen? Weet u van uw collega hoe hij/zij hetzelfde (of een gelijkaardig) leerstofonderdeel zou onderwijzen?
- ✓ Welke gevolgen heeft dit alles voor het al dan niet bestaan van een gemeenschappelijke visie op evaluatie en een gemeenschappelijk evaluatiebeleid? Ontwikkelt u wel eens samen evaluatie-instrumenten? Weet u wat de evaluatiecriteria zijn van uw collega? Bestaan er op schoolniveau richtlijnen die aangeven waaraan een evaluatie zou moeten voldoen?

De noodzaak aan vernieuwingen in het onderwijs is veelal terug te brengen tot veranderde antwoorden op bovenstaande vraag. Pas wanneer deze verandering en de incompatibiliteit van de bestaande instructie- en evaluatiepraktijken om de hedendaagse doelen te bereiken en te evalueren worden gevoeld, zullen leerkrachten een nood ervaren om op zoek te gaan naar andere werkwijzen. De stap van het bepalen van gemeenschappelijke doelen mag dan wel tijdrovend zijn, zonder een algemeen gedragen visie zal een implementatieproces met tegenzin en problemen te kampen hebben.

2 Blijvende gerichtheid op de instructie

Een andere fundamentele hinderpaal richting alternatieve evaluatievormen is de neiging van leerkrachten om zich zozeer op instructieactiviteiten te richten dat de eigenlijke prestaties en resultaten van leerlingen uit het oog worden verloren (Ashbacher, 1993; Wolfe & Miller, 1997). In plaats van zich te focussen op het wenselijke leerlingengedrag neigen leerkrachten ernaar een beroep te doen op klasactiviteiten die ze tijdens hun loopbaan reeds ontwikkelden en succesvol gebruikten. Met andere woorden: leerkrachten zorgen er wel voor dat de nodige inhoudsdomen in hun lessen behandeld, worden maar ze doen dit niet altijd met de prestaties van leerlingen in gedachten. Toch zijn net de gerichtheid op onderwijsdoelen en de directe relatie tussen deze doelen, de instructie en de evaluatie van cruciaal belang om tot goede evaluatievormen te komen. Het geïntegreerde gebruik van evaluatie en instructie maakt dat leerkrachten deels zullen moeten afstappen van de manier waarop ze traditioneel te werk gingen. Om tot een succesvolle implementatie van alternatieve evaluatievormen te komen, moet er een verschuiving plaats vinden van een *activiteitsgeoriënteerde naar een leerdoelgeoriënteerde aanpak*. De vaststelling dat leerkrachten het moeilijk hebben om leerdoelen te expliciteren en om uit te leggen waaruit een goede prestatie bestaat, maakt dan ook dat de invoering van alternatieve evaluatievormen niet zonder slag of stoot gerealiseerd kan worden (Ashbacher et al, 1995).

3 Problemen bij het ontwikkelen en gebruiken van evaluatiecriteria

De aard van alternatieve evaluatievormen vereist dat leerkrachten duidelijke criteria ontwerpen om de beoordeling van de leerlingenprestaties op te baseren. Zowel het ontwikkelen als het gebruiken van evaluatiecriteria blijken echter niet vanzelfsprekend (Ashbacher, 1993). Leerkrachten hebben aanvankelijk de neiging om eerder op zoek te gaan naar taken die geschikt zijn om bepaalde vaardigheden te evalueren dan om evaluatiecriteria te ontwikkelen. Ze ontwijken vragen naar expliciete evaluatiecriteria. De criteria die ze jarenlang gebruikt hebben zijn zodanig geïnternaliseerd dat ze het moeilijk vinden om deze criteria uitdrukkelijk te expliciteren. Toch is een succesvolle invoering van alternatieve evaluatievormen in sterke mate afhankelijk van de duidelijkheid en de openheid waarmee leerkrachten met evaluatiecriteria omspringen. Leerkrachten moeten dan ook voldoende kansen en tijd krijgen om met de achterliggende ideeën vertrouwd te worden.

Bovendien kan het gebruik van evaluatiecriteria leerkrachten aanvankelijk een onbehaaglijk gevoel geven. Ze hebben plotseling de indruk dat ze verantwoording moeten afleggen voor de beoordelingen die ze uitspreken. Terwijl ze vroeger zelden of nooit een bepaalde beoordeling moesten legitimeren, maakt de beschikbaarheid van de evaluatiecriteria dat zowel leerlingen als hun ouders kritisch kunnen toekijken op de beoordelingen. Uit hun beoordelingen zou bijvoorbeeld wel eens kunnen blijken dat leerkrachten onbewust minder consistent zijn in hun beoordelingen dan zichzelf en andere betrokkenen vermoeden.

Dergelijke moeilijkheden in het formuleren van beoordelingen en criteria kunnen omschreven worden als een vorm van 'beoordelingsangst' bij leerkrachten (Ashbacher, 1993). Binnen een assessmentcultuur kunnen leerkrachten de indruk krijgen dat ze zelf beoordeeld worden op basis van de beoordelingen die ze uitspreken. Leerkrachten zijn zich ervan bewust dat beoordelingen zowel voor de leerlingen als voor henzelf gevolgen kunnen hebben. Deze vaststelling kan leerkrachten intimideren waardoor ze het uitspreken van beoordelingen trachten te vermijden. Het gebruik van evaluatiecriteria is dus zeker niet vanzelfsprekend.

4 Tijdsintensief karakter van vernieuwingen

Terwijl bovenstaande valkuilen veelal niet bewust worden beleefd, wordt het tijdsgebrek door leerkrachten zelf als een zeer belangrijke hinderpaal omschreven. De implementatie van alternatieve evaluatievormen moet goed doordacht gebeuren en is dus inderdaad tijdsintensief. Daar komt bij dat scholen verschillende andere vernieuwingen tegelijkertijd trachten in te voeren. Men mag dus niet teveel op een te korte termijn willen realiseren. Dat het niet nemen van voldoende tijd om alternatieve evaluatievormen te implementeren tot problemen kan leiden, blijkt uit volgende elementen. Potentiële valkuilen dus.

- Alternatieve evaluatievormen zijn tijdsintensief omdat ze verwachten dat het gehele evaluatieproces aandachtig doorlopen wordt. Dit betekent uitgebreide aandacht voor de evaluatiedoelen, criteria, standaarden en voor de consistentie van de beoordelingen. Maar ook alvorens effectief aan de slag kan worden gegaan, is er reeds sprake van een aanzienlijke tijdsbesteding. De betrokkenen moeten immers vertrouwd worden met de centrale ideeën en het woordgebruik, er moeten prioriteiten worden vooropgesteld en er moet een consensus worden bereikt over de voorop te stellen doelen. In het bijzonder portfolio-evaluatie vereist een grotere tijdsinvestering van leerlingen en leerkrachten dan andere evaluatiemethoden. Goede portfolioprogramma's kunnen dan ook niet tot stand komen zonder aanzienlijke inspanningen vanwege beleidsvoerders, leerkrachten en leerlingen (OERI, 1993). Van leerkrachten wordt niet alleen een goede kennis van het onderwezen curriculum en van instructiemethodes verwacht; maar ook beschikbare tijd om te plannen, te overleggen met andere leerkrachten, strategieën en materialen te ontwikkelen, bijeen te komen met individuele leerlingen en in kleine groepen en uiteraard ook om de portfolio's te bestuderen en te beoordelen.
- De overgang naar alternatieve evaluatievormen en onderwijsbenaderingen blijkt ook voor leerlingen vaak niet vanzelfsprekend. Leerlingen moeten worden voorbereid op het dragen van meer verantwoordelijkheid en op succesvol reflecteren. De ervaring leert volgens Dochy & Segers (1999) dat leerlingen vaak te kampen hebben met een vorm van onzekerheid. Soms zijn leerlingen gewoonweg te weinig vertrouwd met het geven van open antwoorden (Fisher & King, 1995). Ze zijn het niet gewend zelf verantwoordelijkheden op te nemen, en aan de hand van levensechte taken te bewijzen dat er geleerd werd. Ook een zekere angst voor het evalueren van anderen en het geëvalueerd worden door anderen blijkt leerlingen er van te weerhouden vlot de stap naar het ge-

bruik van bepaalde vormen van alternatieve evaluatie te zetten. Leerkrachten hebben dus de belangrijke taak leerlingen te begeleiden in deze overgangsfase. Ook de leerlingen moeten hiertoe over de nodige tijd kunnen beschikken.

Zowel het komen tot veranderingen in de evaluatiecultuur als het effectief ontwikkelen en gebruiken van alternatieve evaluatievormen vergen een aanzienlijke tijdsinvestering. Met bovenstaande elementen in het achterhoofd lijkt het alsof het invoeren van alternatieve evaluatievormen simpelweg te veel vergt van de reeds drukke bezette leerkrachten. De sleutel tot een succesvolle invoering van alternatieve evaluatievormen bestaat er echter niet in om meer tijd te vinden, maar wel in het herverdelen van de beschikbare tijd met een vernieuwd doel voor ogen. 'Werk slimmer, niet harder' is dus de boodschap (Wiggings, 1998). Eenmaal het gebruik van alternatieve evaluatievormen in voege is, zal er bovendien meer tijd beschikbaar zijn. Rubrieken moeten immers niet steeds opnieuw ontwikkeld worden, leerkrachten gaan samenwerken, de leerlingen worden in de evaluatie betrokken. Maar toch, gezien de aanzienlijke taakbelasting van leerkrachten hoeft het niet te verwonderen dat zij niet onmiddellijk staan te springen om meer complexe, meer uitgewerkte en tijdsintensieve evaluatiemethodes in te voeren. Zoals elke vernieuwing zal het invoeren van alternatieve evaluatievormen aanvankelijk complex en tijdrovend lijken en dus op weerstanden stuiten.

Even stilstaan... 5.2: De noodzaak van tijdsinvestering

Gaat u akkoord met volgende stellingen? Waarom (niet)?

1. Het ontwikkelen en invoeren van alternatieve evaluatievormen leidt tot uiterst interessante kansen om leerkrachten te engageren in de meest fundamentele discussies omtrent wat ze willen bereiken en hoe dat dient te gebeuren.
2. De positieve gevolgen van alternatieve evaluatievormen zijn in grote mate toe te schrijven aan het feit dat leerkrachten deelnemen in het proces van 'consensus building' omtrent waardevolle leeruitkomsten en relevante evaluatiecriteria.
3. Een invoering van alternatieve evaluatievormen is enkel waardevol als deze veel tijd in beslag neemt.

5 Onvoldoende ondersteuning en training van leerkrachten

Leerkrachten die reeds trachtten portfolio's in hun onderwijspraktijk te integreren, geven aan behoefte te hebben aan bijkomende sturing en ondersteuning vanwege collega's of externe experts die reeds ervaring hebben met de betreffende evaluatiemethode. Deze steun mag zich niet beperken tot morele ondersteuning. Klassen die zich organiseren volgens alternatieve evaluatievormen zijn ook ten aanzien van materiële aspecten verschillend van traditionele klaslokalen (flexibiliteit, lawaai, gebruik van ICT). Afwezigheid van een geschikte omkadering kan het implementatieproces aanzienlijk bemoeilijken (Wolfe & Miller, 1997).

Ondersteuning en training zijn niet alleen belangrijk bij het realiseren van een assessmentcultuur en het uit de grond stampen van alternatieve evaluatiebenaderingen. Ook om een kwalita-

tief gebruik van alternatieve evaluatievormen te garanderen, is de nodige ondersteuning onontbeerlijk. Dit geldt bijvoorbeeld wanneer kwaliteitskenmerken zoals validiteit en betrouwbaarheid aan de orde zijn. Concreter: zowel over verschillende beoordelingsmomenten als over verschillende beoordelaars heen moet de consistentie verzekerd worden. Leerkrachten moeten in zulke zaken getraind worden. Wanneer de evaluatie niet valide en betrouwbaar gebeurt, dan zal de legitimiteit van de verkregen gegevens door de leerlingen in twijfel worden getrokken en dan zullen de leerlingen na een tijd de evaluatiemethode in zijn geheel verwerpen. Men moet dus eerst goede beoordelingen kunnen garanderen alvorens leerlingen te beoordelen.

Een gebrek aan ondersteuning op ieder relevant terrein kan de succesvolle implementatie van alternatieve evaluatievormen in het gedrang brengen. Scholing en ondersteuning zullen dan ook een belangrijke plaats innemen wanneer we aandacht besteden aan een aantal mogelijkheden om valkuilen te vermijden. Hierbij moet gestreefd worden naar een complementaire invulling van preservice en inservice training.

6 Problemen bij de leerlingen

Leerlingen die niet de kans krijgen om de nodige vaardigheden te ontwikkelen om succesvol mee te werken aan alternatieve evaluatievormen dreigen tot minder goede prestaties te komen (Ashbacher et al, 1995). Men mag er niet van uitgaan dat leerlingen als vanzelf voldoende voorbereid zullen zijn om succesvol te participeren aan het gebruik van alternatieve evaluatievormen. Ook met betrekking tot leerlingen bestaan er een aantal valkuilen. Dit lijkt bijvoorbeeld logisch wanneer het gaat over de mate waarin leerlingen het Nederlands beheersen. Toch zijn er ook elementen die men op het eerste zicht uit het oog zou kunnen verliezen. Hierbij denken we aan metacognitieve vaardigheden, de gewoonte om volledige en uitgewerkte antwoorden te geven, het inzien van de leerkansen die zelfevaluatie kan bieden, de motivatie van leerlingen en dergelijke. In hoofdstuk twee geven we aan hoe een aantal van deze valkuilen bij leerlingen kunnen vermeden worden.

Leerkrachten moeten de leerlingen helpen om de noodzakelijke vaardigheden te ontwikkelen om met vaardigheidstaken en zelfevaluatie aan de slag te gaan. Het belang van deze stap kan moeilijk overschat worden. Slechte resultaten bij de aanvang van het gebruik van alternatieve evaluatievormen kunnen de gedrevenheid en motivatie bij leerkrachten én leerlingen immers negatief beïnvloeden. Het is dan ook zeer belangrijk een duidelijk zicht te hebben op de voor-kennis van alle leerlingen om het gebruik van alternatieve evaluatiemethoden op een gepaste wijze vorm te geven.

Hoofdstuk 2: Hoe valkuilen vermijden?

Nog belangrijker dan een beschrijving van potentiële valkuilen zijn aandachtspunten die maken dat valkuilen kunnen worden vermeden. In antwoord op de boven beschreven valkuilen willen

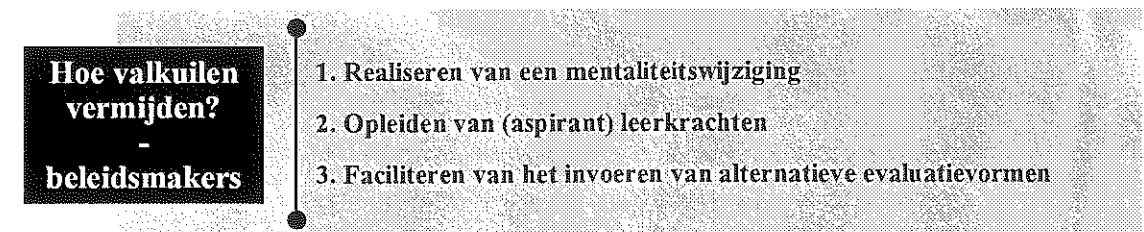
we een aantal elementen beschrijven die een faciliterende werking hebben op de implementatie van alternatieve evaluatiemethodes.

De meest relevante aspecten structureren we op basis van de voornaamste betrokkenen bij het evaluatieproces in het algemeen en bij implementatieproces in het bijzonder, namelijk beleids-makers, leerkrachten, leerlingen en ouders. Voor elke groep rapporteren we een aantal voorwaarden en opdrachten om te komen tot een succesvolle implementatie van alternatieve evaluatievormen.

1 Beleidsmakers

Het begrip beleidsmakers vatten we hier ruim op. We verstaan er alle personen en organen onder die het klasniveau overstijgend meewerken aan het vorm geven aan evaluatie. De rol van beleids mensen situeren wij voornamelijk in het realiseren van de reeds herhaaldelijk benadrukte mentaliteitswijziging, in het scholen van leerkrachten en in het faciliteren van het gebruik van alternatieve evaluatievormen.

Figuur 5.2: De rol van beleidsmakers



1.1 Realiseren van een mentaliteitswijziging

Het onderschrijven door de individuele leerkrachten van de idealen waarop het gebruik van alternatieve evaluatiemethodes is gebaseerd, is één van de belangrijkste facilitatoren van het implementatieproces (Wolfe & Miller, 1997). Toch eindigt het verhaal niet bij de leerkrachten. Ook de weerzin bij andere betrokkenen moet worden overwonnen. Informatie en professionele ontwikkeling moeten zich dus niet beperken tot leerkrachten. Ook beleidsmakers zelf, schoolbesturen, ouders en leerlingen moeten de beoogde mentaliteitswijziging onderschrijven. Om dit te bereiken moeten allen de noodzaak van alternatieve evaluatievormen aanvoelen. Hiertoe is het belangrijk de doelen en de verschillende mogelijkheden van alternatieve evaluatievormen uit te leggen en te illustreren. Om tot een mentaliteitswijziging te komen, is er dus een zeer belangrij-

ke rol weggelegd voor goede voorbeelden. De betrokkenen moeten de meerwaarde van alternatieve evaluatievormen ervaren. Dit kan door de betrokkenen (bijvoorbeeld ouders) te betrekken in het evaluatieproces (bijvoorbeeld in het vooropstellen van doelen of in het beoordelen van leerlingenprestaties). Hetzelfde geldt voor leerkrachten. Ook zij zullen sneller geneigd zijn tijd in vernieuwingen te investeren indien ze effectief ervaren dat deze bijdragen tot betere leerprocessen van leerlingen. Communicatie en het uitwisselen van ervaringen spelen dan ook een belangrijke rol in het realiseren van een mentaliteitsverandering bij de verschillende betrokkenen.

1.2 Opleiden van (aspirant) leerkrachten

Leerkrachten vormen de kern van de kanteling naar een assessmentcultuur. Vernieuwingen staan of vallen in belangrijke mate bij de motivatie en de vaardigheden van leerkrachten. Wanneer leerkrachten niet overtuigd zijn van de bruikbaarheid van de alternatieve evaluatievormen en de evaluatiegegevens niet kunnen aanwenden om het leerproces van hun leerlingen te optimaliseren, bestaat het gevaar dat er zich een zekere weerstand ontwikkelt (Arter & Bond, 1996). In het scholen van leerkrachten onderscheiden we daartoe verschillende componenten: een mentaliteitsgerichte, een praktische en een inhoudelijke component. Elke component moet zowel aan bod komen in de initiële lerarenopleiding als in de nascholing waarop leerkrachten een beroep kunnen doen.

Even stilstaan... 5.3: De scholing van leerkrachten

1. Hoe zou u als beleidsmaker uw leerkrachten professionaliseren met het oog op een succesvol gebruik van alternatieve evaluatievormen?
2. Vergelijk uw voorstel tot scholing met de scholing die leerkrachten in het 'Vermont portfolio assessment program' ontvingen (Aschbacher et al., 1995). Het gebruik van portfolio's werd daar voor bepaalde vakken verplicht gemaakt. Welke aspecten lijken u interessant in deze aanpak van de scholing?

Vermont creëerde kansen voor leerkrachten om zich te verdiepen in het portfolioprogramma en om zich voor te bereiden op de implementatie. Tijdens het pilootjaar (1990-1991) sponsorde de overheid hiertoe een centraal georganiseerde bijeenkomst en een aantal regionale workshops. In het volgende schooljaar werd de inservice training van leerkrachten nog uitgebreid en werden workshops ingericht die leerkrachten voorbereiden op het scoren en beoordelen van leerlingenprestaties. Alle leerkrachten ontvingen een 'bronnenboek' met bruikbare voorbeelden en een 'leerkrachtengids' met praktische richtlijnen. Eenmaal leerkrachten met de instrumenten aan de slag waren, werden regionale netwerken opgericht om in bijkomende training te voorzien en om aan expertisedeling te doen.

De eerste component die in de scholing moet worden aangepakt, is de **mentaliteitsgerichte component**. Leerkrachten moeten overtuigd worden van de meerwaarde van alternatieve evaluatievormen en moeten de tekorten van (hun eigen) traditionele evaluatiepraktijk leren inzien. Op die manier worden leerkrachten medebepleiters van de assessmentcultuur en van het gebruik van alternatieve evaluatievormen. Pas wanneer leerkrachten een duidelijk zicht hebben op wat

de doelen zijn van het gebruik van andere evaluatievormen en ze met een duidelijke visie aan de slag kunnen, heeft het zin om met de praktische component van de scholing aan te vangen.

Ook al is evaluatie een frequent gebeuren in de Vlaamse klaslokalen, leerkrachten hebben zelden de kans om de noodzakelijke vaardigheden – zoals het ontwikkelen van evaluatiecriteria en het gebruik van rubriecken – te leren en te oefenen. Een aandachtspunt dat herhaaldelijk wordt benadrukt is dan ook de noodzaak om leerkrachten adequaat te trainen in het gebruik van alternatieve evaluatiemethoden. Vaardigheden en know-how zijn cruciaal. Hierin schuilt de **praktische component** van de scholing. Beleidsmakers moeten leerkrachten voorbeelden aanbieden én de mogelijkheden bieden om zichzelf te scholen. Zij moeten zo een klimaat scheppen waarin leerkrachten gestimuleerd worden om zichzelf te vervolmaken. Wanneer een dergelijk constructief klimaat niet kan gerealiseerd worden, dan kan de succesvolle implementatie van alternatieve evaluatievormen op de helling komen te staan. Vandaar het belang van de combinatie van het ter beschikking stellen van voorbeeldtaken en de ondersteuning van leerkrachten bij het ontwikkelen van eigen taken. Wanneer leerkrachten problemen ervaren bij het vinden van geschikte taken en hiermee nergens terecht kunnen, dan zullen zij teruggrijpen naar oude taken. Dit resulteert dan bijvoorbeeld in portfolio's waarin uiterst traditionele drill-oefeningen zijn opgenomen.

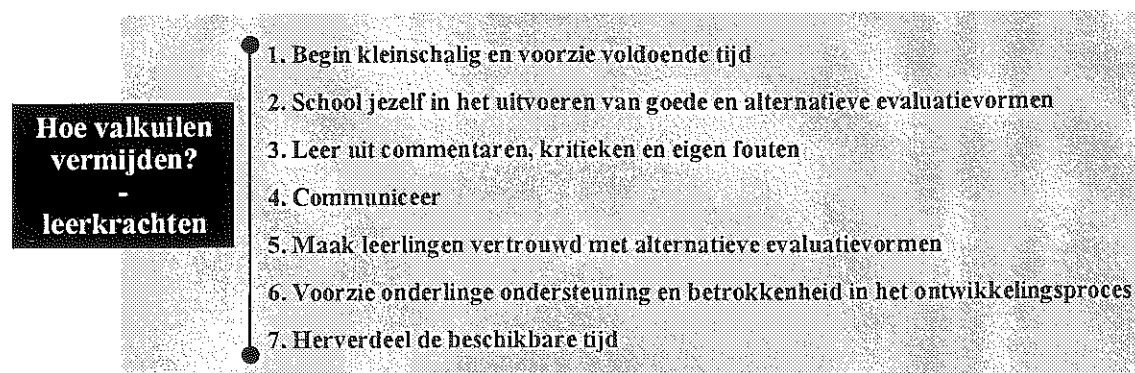
Niet alleen de kennis en vaardigheden om te evalueren, moeten de nodige aandacht krijgen. De nadruk die op het ontwikkelen van criteria en standaarden wordt gelegd, impliceert dat leerkrachten ook expert moeten zijn in verschillende inhoudsdomeninen. De scholing van leerkrachten heeft dus ook een **(vak)inhoudelijke component**. Dit geldt bijvoorbeeld wanneer moet nagegaan worden welke aspecten kunnen onderscheiden worden in complexe vaardigheden binnen hun vakdomein. Het zou onterecht zijn er van uit te gaan dat leerkrachten voor alle vakken die ze geven telkens over de noodzakelijke achtergronden beschikken. Om goede evaluatietaken en beoordelingsinstrumenten te ontwikkelen, moeten leerkrachten dus ook voldoende inhoudelijk geschoold zijn en blijven.

1.3 Faciliteren van het invoeren van alternatieve evaluatievormen

Zowel het invoeren als het uiteindelijke gebruik van alternatieve evaluatievormen zijn veeleisend wat betreft de inzet van menselijke en financiële middelen binnen scholen. Het is dan ook onrealistisch en onwenselijk te denken dat grote vernieuwingen zoals het gebruik van alternatieve evaluatievormen enkel door vrijwilligers kunnen gerealiseerd worden. Wanneer het de overheid menens is wat betreft het promoten van alternatieve evaluatievormen, dan zal zij het gebruik ervan op talrijke wijzen moeten faciliteren. Vernieuwingsprocessen creëren verscheidene behoeftes waaraan de overheid zoveel mogelijk moet tegemoetkomen: tijd, geld, ideeën en expertise. De overheid en schoolbesturen moeten het bijvoorbeeld mogelijk maken dat de nodige leerkrachten tijdelijk vrijgesteld worden van lesopdrachten om binnen de eigen school te werken aan de noodzakelijke houdingen, doelen en structuren (Westerberg, 1996). Hiertoe moeten op de verschillende bestuursniveaus de nodige fondsen worden voorzien. De facilite-

rende rol beperkt zich echter niet tot binnenschoolse activiteiten. Het is niet de bedoeling dat elke school en elke leerkracht het warm water gaat uitvinden. Vandaar de belangrijke taak voor de overheid om overkoepelende activiteiten op te zetten waarin voor -verschillende inhoudsdo-
meinen en verschillende groepen- evaluatietaken en beoordelingsinstrumenten worden ontwik-
keld en aan de scholen ter beschikking gesteld.

De rol van beleidsmakers en het belang van het meso-niveau in de implementatie van alterna-
tieve evaluatievormen kunnen moeilijk overschat worden. De visie van de school, de aard van
de professionele relaties tussen leerkrachten, de ondersteuning die leerkrachten kunnen genieten
en de mate van samenwerking, zijn sterk afhankelijk van het schoolbeleid. In veel van de aan-
dachtspunten die we presenteren spelen het schoolbeleid en de schoolleiding een belangrijke
rol.



2 Leerkrachten

Beleidsmakers kunnen enkel een gepaste context creëren waarin het gebruik van alternatieve
evaluatievormen mogelijk wordt. Het is aan leerkrachten om in die context aan de slag te gaan.
Hierbij kunnen ze volgende zeven raadgevingen in acht nemen.

Figuur 5.3: Aandachtspunten voor leerkrachten

2.1 Begin kleinschalig en voorzie voldoende tijd

Om aan te vangen met alternatieve evaluatievormen is het aangewezen om te beginnen met het
uitwerken van een eerder beperkt curriculum domein. Voor dat domein kunnen doelen voorop-
gesteld worden waarvoor vervolgens geschikte activiteiten en beoordelingsinstrumenten ont-
wikkeld worden. Op die manier beschik je als leerkracht over een beheersbare opdracht zonder
te worden overweldigd en het overzicht kwijt te geraken. Eenmaal men vertrouwd is met het
evaluatieproces en met de alternatieve evaluatievormen kan het gebruik ervan worden uitge-
breid. Dit betekent bijvoorbeeld dat een kerngroep van leerkrachten voor zichzelf een 2- of 3-
jarenplan kan uitwerken waarin ze zich eerst de concepten van alternatieve evaluatievormen
eigen maken alvorens het gebruik ervan systematisch te verruimen.

Een te snelle invoering van alternatieve evaluatievormen moet dus ten alle tijden vermeden
worden. Pas wanneer leerkrachten en leerlingen zich de doelen van alternatieve evaluatievor-
men eigen hebben gemaakt, wanneer ze zich bewust zijn van de kwaliteitseisen inzake evaluatie
en wanneer ze de praktische en inhoudelijke vaardigheden hebben om vaardigheidstaken te
ontwerpen en te beoordelen, moet aan de systematische uitbreiding van het gebruik van alterna-
tieve evaluatievormen gedacht worden.

Even stilstaan... 5.4: Tijdsplanning van een invoeringsproces

Westerberg – een directeur die vaardigheidsevaluatie in zijn school implementeerde – stelt: “Every
thing we have done has taken longer than we thought it would” (Westerberg, 1996, p.2).

1. Sluit dit aan bij uw eigen ervaringen met veranderingsprocessen?
2. Wat kunnen we hieruit leren voor de implementatie van alternatieve evaluatievormen?

Mensen die reeds ervaring hebben met veranderingsprocessen zijn zich bewust van het feit dat
tijdsplanningen die leerkrachten en scholen vooropstellen om een bepaald traject af te leggen
veelal weinig realistisch zijn. Bij het begeleiden van scholen wordt er door sommigen bijvoor-
beeld van uitgegaan dat in de oorspronkelijke planning de hoeveelheid beschikbare tijd moet
verdubbeld worden. Anderen gebruiken dan weer een vuistregel die er op neerkomt dat de oor-
spronkelijke planning met pi ($\pi = 3.14$) moet vermenigvuldigd worden om tot een haalbare en
draagbare planning te komen. Bij het invoeren van alternatieve evaluatievormen moet er dus
een sterke realiteitszin aan de dag worden gelegd. Het is immers vanzelfsprekend dat naïviteit
tijdens de planning kan leiden tot onnodige frustraties en teleurstellingen wanneer eerder voor-
opgestelde doelen en deadlines onhaalbaar blijken.

2.2 School jezelf in het uitvoeren van goede en alternatieve evaluatievormen

Het is uiteraard niet alleen de taak van beleidsmakers om leerkrachten te scholen. Leerkrachten
moeten zich -om goede evaluaties te kunnen verzekeren- ook op eigen initiatief vervolmaken.
De vraag welke evaluatie-aspecten hiertoe in aanmerking komen wordt best door de leerkrach-
ten zelf beantwoord (dit kan bijvoorbeeld aan de hand van door de overheid ter beschikking
gestelde beoordelingsinstrumenten om de evaluatie zelf onder de loop te nemen). Terwijl de ene

leerkracht problemen ondervindt bij het vinden van geschikte evaluatietaken, kan de ander meer nood hebben aan hulp bij het vooropstellen van criteria en standaarden. De beschikbare nascholing mag dus niet enkel aanbodgericht zijn; leerkrachten hebben de verantwoordelijkheid om hun eigen nascholingsbehoeftes te detecteren, te verwoorden en vraaggericht te vervullen.

Even stilstaan... 5.5: Uw eigen nascholingsbehoefte

Het goed gebruik van evaluatievormen betekent dat de evaluatie valide, veralgemeenbaar en rechtvaardig is. Het betekent verder ook dat de evaluatie tegemoet komt aan de specifieke doelen die men beoogt en aan specifieke kenmerken van de leerlingen die van invloed kunnen zijn op de prestatiegegevens. Deze greep uit de vereisten van een goed opgezette evaluatie illustreert de complexe en technische vaardigheden waarover een leerkracht moet beschikken.

1. Voor welke aspecten van het gebruik van alternatieve evaluatievormen heeft u de meeste behoefte aan bijkomende scholing? Met andere woorden: waarin vindt u dat uzelf nog kan verbeteren om succesvol gebruik te kunnen maken van alternatieve evaluatievormen?
2. Denkt u dat uw collega's dezelfde of gelijkaardige nascholingsbehoeftes hebben?

2.3 Leer uit commentaren, kritieken en eigen fouten

Veranderingen en vernieuwingen leiden tot spanningen en kritieken. De manier waarop een schoolteam hiermee omspringt, is bepalend voor het welslagen van de vernieuwing. Spanningen en kritieken mogen niet uit de weg worden gegaan, zelfs indien ze niet constructief en goed bedoeld zijn. Kritieken op het proces en het doel kunnen immers aangewend worden om het proces en het doel te verbeteren. Westerberg (1996) concludeert uit eigen ervaringen dat het aangewezen is om kritieken te behandelen als gelijk welk ander argument en om ze gezamenlijk te onderzoeken. Kritieken kunnen immers legitieme argumenten zijn. Bovendien kan een constructief scepticisme problemen aan het licht brengen alvorens ze het invoeringsproces in gevaar dreigen te brengen.

Een gelijkaardige redenering geldt voor het omgaan met eigen mislukkingen en fouten. Leerkrachten moeten elkaar de kans geven om in een veilige context te experimenteren en op zoek te gaan. Pas wanneer openlijk gecommuniceerd wordt over de eigen fouten, kan vermeden worden dat fouten zich herhalen en een eigen leven gaan leiden. Vergeet niet dat uit zorgvuldige analyse van de oorzaken van een bepaalde fout vaak meer kan geleerd worden dan uit het laten passeren van een op het eerste zicht goed voorbeeld.

2.4 Communiceer

Over vernieuwingen moet zowel intern als extern gecommuniceerd worden. In de eerste plaats om iedereen te informeren over op til staande activiteiten en veranderingen. Het kan bijvoorbeeld niet zijn dat ouders pas op een schoolrapport merken dat het beoordelingssysteem en de manier van rapporteren over evaluatieresultaten grondig hervormd zijn. Door te communiceren

creëert men duidelijkheid. Dit betekent dat misverstanden worden vermeden, dat toegelicht wordt waar men naar toe wil en dat verantwoord wordt waarom bepaalde zaken niet meer zijn zoals ze waren. Communiceren kan bijvoorbeeld aan de hand van korte nieuwsberichten (in verslagen of in het schoolkrantje). Niets garandeert echter dat die zaken ook gelezen worden. Vandaar het belang van mondelinge communicatie in formele en informele contexten. Op die manier wordt het eenrichtingsverkeer doorbroken en kunnen vragen, onduidelijkheden, bekommernissen maar ook alternatieven en relevante argumenten aan het licht komen.

Communicatie maakt leden deel van een groep. Het ter beschikking hebben van een groep waarmee men kan samenkomen en ideeën uitwisselen, blijkt een cruciale factor in het succesvol implementeren van alternatieve evaluatievormen. Leerkrachten vinden het samen werken aan vernieuwing plezieriger en minder bedreigend dan alleen aan de slag gaan (Ashbacher, 1993).

Om te kunnen communiceren, moet een gemeenschappelijke taal worden gesproken. Wat alternatieve evaluatievormen betreft, is dit zeker een aandachtspunt. De vele termen worden in de literatuur en in de praktijk immers op verschillende manieren ingevuld. Om de interne en externe communicatie succesvol te laten verlopen, is er dus nood aan conceptuele duidelijkheid. Deze kan er bijvoorbeeld in bestaan dat vroeg in het invoeringsproces een begrippenlijst wordt samengesteld waarin elk van de gebruikte termen voor gebruik in de eigen school wordt omschreven.

2.5 Maak leerlingen vertrouwd met alternatieve evaluatievormen

Men moet steeds voor ogen houden dat leerlingen voldoende tijd nodig hebben om zich aan de alternatieve evaluatiemethoden aan te passen. Leerlingen moeten zichzelf als doelmatig beleven en gemotiveerd zijn om inspanningen te leveren en steeds beter te presteren. Het gebruik van alternatieve evaluatievormen impliceert echter niet automatisch dat leerlingen zullen meestappen in het traject dat men wil uitstippelen. Bij wijze van voorbeeld: bij zelfevaluatie verwacht men van leerlingen dat ze reflecteren en denkstrategieën ontwikkelen. Dergelijke vaardigheden moeten echter uitdrukkelijk onderwezen worden. Dit geldt bijvoorbeeld voor de criteria en standaarden die gehanteerd zullen worden bij de beoordeling van de evaluatietaken en die leerlingen dus moeten leren gebruiken. Deze vaardigheden kunnen doorheen het instructieproces eerst ontwikkeld worden. Hieruit blijkt de link tussen de visies op onderwijs, onderwijsdoelstellingen, curriculum en evaluatie.

Om leerlingen vertrouwd te maken met en te motiveren voor het gebruik van alternatieve evaluatievormen kunnen we een aantal richtlijnen naar voren schuiven (Ashbacher et al, 1995; Dochy & Segers, 1999). De eerste bijeenkomsten met de leerlingen kunnen gebruikt worden om de onderwijsaanpak en de evaluatievormen toe te lichten. Men kan op dat moment concrete voorbeelden van de evaluatievorm geven en bespreken. Het is eveneens zinvol de criteria bij de komende evaluatie te bespreken en/of samen te ontwikkelen. Leerkrachten moeten alvorens de instructie van een leerstofonderdeel aanvangt de beschikbare rubrieken en voorbeelden van

goede prestaties met de leerlingen delen. Het mag dus zeker niet zo zijn dat leerlingen moeten gissen naar wat goede prestaties zijn. Wanneer leerlingen niet weten wat ze fout doen, hoe kunnen ze zichzelf dan verbeteren? Tijdens de onderwijsbijeenkomsten moet ook regelmatig benadrukt worden dat niet alleen reproductie van kennis geëvalueerd wordt. Tot slot, indien er tussentijdse evaluatiegegevens ter beschikking zijn, dan is het aangewezen de leerlingen feedback te geven omtrent die vaardigheden die van tel zijn om te participeren aan alternatieve evaluatievormen (zie ook paragraaf 3).

2.6 Voorzie onderlinge ondersteuning en betrokkenheid in het ontwikkelingsproces

Een andere faciliterende factor is de ondersteuning die leerkrachten kunnen genieten. Leerkrachten geven immers zelf aan dat ondersteuning door anderen (collega-leerkrachten, beleidsmensen, technische experts) een essentiële voorwaarde is om de implementatie van alternatieve evaluatie succesvol te doen verlopen. Samenwerken is dus de boodschap. Eerder dan een inhoud en format opgelegd te krijgen, willen leerkrachten hulp van andere leerkrachten en technische ondersteuning bij het ontwikkelen van eigen evaluatie-instrumenten. Op die manier kan verzekerd worden dat de evaluatie zeker aangepast is en bruikbaar zal zijn in hun specifieke onderwijssituatie.

Een interessante mogelijkheid bestaat erin de betrokken leerkrachten samen te brengen in werkgroepen. In dergelijke groepen kunnen zij (bijvoorbeeld maandelijks) ervaringen uitwisselen en publicaties over evaluatie bespreken. Zowel ervaringen omtrent het ontwerpen van evaluatietaken en beoordelingsinstrumenten als ervaringen omtrent strategieën om dergelijke ontwikkelingsprocessen te organiseren zijn uitermate leerrijk en stimulerend. Indien nodig kunnen in de werkgroepen inhoudelijke experts worden uitgenodigd om het eigen blikveld te verstevigen en te verruimen.

De betrokkenheid van leerkrachten is cruciaal in vernieuwingsprocessen. Leerkrachten die het gebruik van alternatieve evaluatievormen genegen zijn, zullen bereid zijn meer te investeren doorheen het invoeringsproces, zullen veranderingen sneller incorporeren in hun klaspraktijk, zullen leerlingen en collega-leerkrachten inspireren en motiveren, en zullen meer volharden wanneer ze geconfronteerd worden met tegenwerking (Ashbacher, 1993).

2.7 Herverdeel de beschikbare tijd

De nodige tijd vinden blijkt één van de meest gemaakte bedenkingen van leerkrachten bij het implementatieproces van alternatieve evaluatievormen. De extra taken c.q. tijd die leerkrachten moeten besteden aan evaluatie worden gezien als een bijkomende taakbelasting bovenop de andere taken. Als leerkrachten evaluatie blijven zien als een afzonderlijke fase in het didactische proces kan tijdsgebrek inderdaad een hinderpaal vormen. Net het integreren van evaluatie en instructie maakt echter dat er tijd kan vrijgemaakt worden. Op die manier hoeft het gebruik van

alternatieve evaluatievormen niet meer tijd te vergen van leerkrachten dan voorheen. Alternatieve evaluatie neemt immers de plaats in van andere onderwijskundige en administratieve activiteiten. In plaats van de volledige dag les te geven en op het einde van de dag nog huiswerk en toetsen te moeten verbeteren, organiseert de leerkracht in een assessmentcultuur de evaluatie met de leerlingen tijdens de instructie. Dit betekent bijvoorbeeld dat leertaken tevens aangewend worden als evaluatietaken en vice versa, dat het ontwikkelen van criteria ter beoordeling van de leerlingenprestaties eveneens wordt aangezien als een belangrijk leermoment en dat leerlingen graag samenwerken en zeer veel kunnen leren van hun klasgenoten (Fisher & King, 1995). Wanneer de leerlingen actief aan het werk zijn, beschikt de leerkracht over de nodige tijd om te observeren, om voorbeelden van prestaties van leerlingen te verzamelen, en om samen met de leerlingen te evalueren. Op die manier wordt evaluatie een gezamenlijke verantwoordelijkheid van de leerkracht en de leerlingen en een integraal onderdeel van het lesgebeuren. Wat op het eerste zicht tijdrovend lijkt te zijn, zou dus wel eens bijkomende tijd kunnen vrijmaken voor bijvoorbeeld bijkomende begeleiding van en gerichtere feedback voor de leerlingen. Het is niet zozeer het gebruik van alternatieve evaluatie dat tijdrovend is, het is het ontwikkelen van taken en beoordelingsinstrumenten dat extra eisen stelt aan leerkrachten. Net hierin mogen leerkrachten echter niet alleen komen te staan.

Of bovenstaand optimisme terecht is, kan momenteel echter nog niet met zekerheid vastgesteld worden. Daarbij denken we aan ervaringen met 'zelfstandig leren' van leerlingen waarbij men ook zou verwachten dat leerkrachten daardoor meer tijd zouden hebben. Toch blijkt dit in realiteit niet altijd het geval.

3 Leerlingen

Men mag van leerlingen niet verwachten dat ze zomaar de stap kunnen zetten van een traditionele onderwijs- en evaluatiecultuur naar een assessmentcultuur. Hiertoe moeten ook zij over aangepaste opvattingen en de nodige vaardigheden beschikken. Ook leerlingen moeten de cultuurswijziging onderschrijven en de meerwaarde van alternatieve evaluatievormen ervaren alvorens zij gedreven in een invoeringsproces zullen meestappen. De motivatie van leerlingen is immers van groot belang om het welslagen van alternatieve evaluatievormen te garanderen. Leerkrachten moeten de leerlingen ondersteunen om de noodzakelijke vaardigheden te ontwikkelen om met vaardigheidstaken en zelfevaluatie aan de slag te gaan. De aard van de vragen, taken en opdrachten bij alternatieve evaluatievormen verwachten van leerlingen immers dat ze complexe denkvaardigheden aanwenden. Alvorens met alternatieve evaluatievormen aan de slag te gaan, is het daarom zeer belangrijk een goed beeld te hebben van de voorkennis van leerlingen ten aanzien van deze vaardigheden (Fisher & King, 1995). Indien nodig moet daarom tijdens de instructie eerst bijkomende aandacht worden besteed aan volgende elementen.

Leerlingen moeten leren hoe te denken en hoe te antwoorden. Daarom moet (vakoverschrijdend) aandacht worden besteed aan de betekenis van termen en activiteiten zoals definiëren, opsommen, vergelijken, samenvatten, bekritisieren, analyseren en evalueren. Dergelijke activiteiten moeten aangeleerd en geoefend worden. Leerlingen moeten de sleutelementen in een

opdracht herkennen en er gepast op kunnen reageren. De aard van de antwoorden en de bijbehorende kwaliteitscriteria bij alternatieve evaluatievormen zijn immers verschillend van traditionele antwoordvormen. Om dit te oefenen kunnen leerlingen bijvoorbeeld aan de hand van duidelijke evaluatiecriteria geanonimiseerde antwoorden van medeleerlingen beoordelen. Om goede antwoorden te leren formuleren, kunnen leerlingen gevraagd worden hun antwoorden te richten aan iemand die een volkomen leek is in het betreffende leerdomein. Leerlingen vinden het bovendien leerrijk om over hun eigen oplossingsprocessen en antwoorden te praten.

De leerlingenbetrokkenheid bij alternatieve evaluatievormen vergt eveneens een zekere kennis van leerlingen. In feite komt het er in de eerste plaats op neer om leerlingen op een gepaste manier te leren omgaan met evaluatiecriteria en standaarden. Op die manier kunnen ze geleidelijk aan hun eigen prestaties leren beoordelen. Merk op dat het ontwikkelen van de vaardigheden van leerlingen om te reflecteren over hun eigen prestaties en leerproces zowel een voorwaarde als een doel zijn van alternatieve evaluatievormen. Een zekere bekwaamheid om (zichzelf) te evalueren is een noodzakelijke eerste stap in het gebruik van alternatieve evaluatievormen. Het openlijk uitwisselen van ervaringen en klassikaal of in groep bespreken en beoordelen van concrete leerlingenprestaties is in die zin zeer leerrijk. Het komt er in dergelijke activiteiten op aan om op zoek te gaan naar kwaliteitskenmerken van deze prestaties en om aan de hand van criteria de kwaliteitsverschillen te expliciteren. In het verdere verloop van de assessmentcultuur moeten deze vaardigheden ook bij een individuele reflectie op de eigen prestaties blijvend ondersteund en ontwikkeld worden. Het is dus aan te bevelen om eerst ervaring op te doen met het evalueren van prestaties van onbekenden en om pas daarna aanzelf- en peerevaluatie te doen. Wanneer leerlingen het gebruik van criteria onder de knie hebben, kunnen zelf-evaluatietechnieken zoals zelf-evaluatielijsten en logboeken in gebruik worden genomen.

4 Ouders

Doorgaans hebben ouders wel een aantal kritieken op de school van hun kinderen maar over de ganse lijn zijn ze veelal tevreden. Zij zijn dan ook niet altijd vanzelfsprekend enthousiast over onderwijsvernieuwingen. Integendeel, wanneer ze geen problemen ervaren, zijn ze eerder voor een behoud van de bestaande aanpak. Uiteraard zijn ook ouders over het algemeen steeds te vinden voor aanzetten tot kwaliteitsverbetering, maar dan wel voor de leerlingen in het volgende schooljaar. Westerberg (1996) stelt: iedereen wil verandering, tot je begint te veranderen. Grof verwoord: ouders willen zekerheid voor hun eigen kind en zullen experimenten liever laten uitstellen tot een volgend schooljaar. Van deze ingesteldheid moet een schoolteam zich zeker bewust zijn. Het is immers noodzakelijk dat ook de steun van ouders voor het gebruik van alternatieve evaluatievormen gewonnen wordt. Net als leerkrachten en leerlingen moeten ouders de mentaliteitswijziging onderschrijven. Ook zij moeten waar nodig overtuigd en bijgeschoold worden. De rol van ouders in vernieuwingsprocessen mag daarom niet onderschat worden.

Bij het invoeren van alternatieve evaluatievormen is het raadzaam ouders te beschouwen als partners van leerkrachten en leerlingen. Dit betekent dat zij betrokken worden in alle aspecten

van het implementatieproces en dat ze aangezien worden als een potentieel rijke bron van informatie. De rol van ouders kan daartoe op verschillende manieren ingevuld worden. We onderscheiden een passieve en een actieve invulling.

Ouders hebben een *passieve* rol in die zin dat ze geïnformeerd moeten worden. Ze moeten ingelicht worden over de inhoud van alternatieve evaluatievormen, over wat er verandert ten opzichte van vroeger, over hoe leerkrachten komen tot beoordelingen over leerlingenprestaties en over de manier waarop over de prestaties van leerlingen gerapporteerd wordt. Hiertoe kan men ouders voorbeelden geven van de gebruikte evaluatietaken en van de gebruikte evaluatiecriteria en rubrieken. Op die manier kunnen zij vertrouwd gemaakt worden met evaluatie in een assessmentcultuur. Hierbij moet voldoende mogelijkheid gecreëerd worden om in te gaan op vragen en om een open discussie te hebben.

Ouders kunnen echter ook op een *meeractieve* manier bij het invoeringsproces worden betrokken. Ze kunnen een belangrijke rol spelen in het communiceren van de noodzaak van vernieuwingen naar andere ouders. Onzekerheid en angst verdwijnen immers wanneer leden van een eigen groep zich met autoriteit uitspreken in het voordeel van een vernieuwing. Dit betekent dat een aantal ouders op een actieve manier wordt ingeschakeld om de passieve rol van ouders in te vullen. De bijdrage van ouders aan het implementatieproces kan echter nog op andere manieren ingevuld worden. Hierbij denken we bijvoorbeeld aan het betrekken van ouders in het vooropstellen van leerdoelen en aan de rol die ouders kunnen spelen in het opsporen van sterke en zwakke punten van hun kinderen. Dergelijke activiteiten maken dat ouders aan den lijve de veranderingen en de potentiële meerwaarde van de vernieuwing ondervinden. Dit kan het implementatieproces enkel ten goede komen.

Samenvatting

Vernieuwingen zoals het invoeren van alternatieve evaluatievormen vergen een grote tijdsinvestering. Tijdsgebrek wordt door leerkrachten zelf dan ook omschreven als de belangrijkste hinderpaal bij het invoeren van alternatieve evaluatievormen. Daarnaast ervaren ze een moeilijk te overwinnen gevoel van onzekerheid. Dit gevoel resulteert uit de vaststelling dat het gebruik van alternatieve evaluatievormen geen beperkte aanpassing is, maar een volledig nieuwe aanpak van onderwijzen vergt (Ashbacher, 1993). Uit bovenstaande beschrijving blijkt dat dergelijke elementen kunnen leiden tot een zekere terughoudendheid bij leerkrachten om met alternatieve evaluatievormen aan de slag te gaan. De implementatie wordt echter eveneens bemoeilijkt door een aantal problemen waarvan leerkrachten zich doorgaans niet echt bewust zijn. Hierbij denken we aan een blijvende gerichtheid op de instructie van bepaalde leerstofonderdelen ten nadele van een focus op de na te streven leerdoelen en aan hun terughoudendheid bij het ontwikkelen en gebruiken van evaluatiecriteria. In het vorige hebben we dergelijke minder vanzelfsprekende valkuilen onder de aandacht willen brengen.

In hoofdstuk 2 hebben we aangetoond dat valkuilen niet onvermijdelijk zijn. Hierbij werd sterk benadrukt dat een succesvolle implementatie van alternatieve evaluatievormen de betrokkenheid en het engagement vergt van zowel beleidsmakers als schoolleiders, leerkrachten, leerlingen en ouders.

Verklarende woordenlijst

Alternatieve evaluatie. Een evaluatiemethode kan als *alternatief* kan beschouwd worden indien ze in voldoende mate aan een aantal eisen voldoet. Deze zijn: aangepastheid, constructiegerichtheid, levensechtheid, voldoende leerlingenbetrokkenheid en de integratie van evaluatie en instructie.

Assessment. Assessment is het proces dat erop gericht is te ontdekken wat een persoon kent en kan. Bij assessment is men voornamelijk begaan met de individuele persoon, de student of de leerling.

Assessmentcultuur. In de assessmentcultuur wordt de integratie tussen instructie en assessment sterk benadrukt. De leerling is een actieve participant die mede verantwoordelijkheid draagt voor het evaluatieproces, die zichzelf evalueert, die reflecteert en samenwerkt in een aanhoudende dialoog met de leerkracht. De evaluatie zelf kan verschillende vormen aannemen. De taken zijn interessant, betekenisvol, authentiek, uitdagend en vergen telkens een vorm van onderzoek. De leerlingen participeren in de ontwikkeling van criteria en standaarden die zullen gehanteerd worden om hun prestaties te beoordelen. Zowel het proces als het product worden geëvalueerd en gerapporteerd op een manier die kan variëren van een allesomvattend cijfer tot het opstellen van een profiel of uitvoerige beschrijving.

Authenticiteit - Levensechtheid. De levensechtheid (of authenticiteit) van een evaluatie verwijst naar de eigenschap dat een evaluatieopdracht of -taak die de leerlingen/studenten moeten vervullen, op zich als betekenisvol en waardevol wordt ervaren.

Authentieke evaluatie beoordeelt de mogelijkheden van leerlingen aan de hand van realistische taken en situaties. Authentieke evaluatietaken verwachten van leerlingen dat ze vaardigheden en competenties aantonen op een manier waarop professionals die vaardigheden en competenties in de dagelijkse realiteit aanwenden. We spreken dus van authentieke evaluatie wanneer de evaluatietaken gelijken op de taken die professionals ondernemen of -ruimer opgevat- wanneer de taken authentiek zijn in die zin dat ze de zogenaamde realistische wereld buiten de school representeren.

Betrouwbaarheid. De betrouwbaarheid is de mate waarin men staat kan maken op de evaluatieresultaten. Dat wil zeggen de mate waarin de beoordelingen consistent, nauwkeurig en reproduceerbaar zijn, kortom vrij van meetfouten. Bij een betrouwbare evaluatie mag het evaluatiereultaat niet beïnvloed worden door storende factoren.

Checklists. Checklists (of observatieschalen) zijn een soort evaluatielijst. Bij het gebruik ervan beantwoordt men enkel de vraag of een gedraging zich al dan niet heeft voorgedaan, zonder te beoordelen hoe goed elk van de gedragingen uitgevoerd werd. De schaal beperkt zich tot de beoordelingen zoals ja - nee, aanwezig - afwezig, volbracht - niet volbracht.

Collaboratieve evaluatie. Collaboratieve evaluatie (of co-evaluatie) is een tussenstadium tussen de traditionele eenzijdige evaluatie door de leerkracht enerzijds en zelf- en peerevaluatie anderzijds. In een collaboratieve evaluatiemethode komen de leerling, de medeleerlingen en de leerkracht samen om te evalueren. Collaboratieve evaluatie creëert aldus een dialoog tussen de leerkracht en de leerlingen tijdens het beoordelen. De leerling is niet de eindverantwoordelijke voor de evaluatie maar werkt wel mee aan het evaluatieproces.

Constructiegerichtheid. Constructiegerichtheid is het essentiële kenmerk van vaardigheidsevaluatie. Het kan kort en krachtig als volgt omschreven worden: leerlingen moeten zelf antwoorden en/of producten construeren. Men richt zich niet zozeer op de reproductie van kennis, maar veeleer op het evalueren van vaardigheden en attitudes. De vraag is of de lerende in staat is om de (voor)kennis te gebruiken bij het oplossen van nieuwe problemen.

Criteria. Criteria zijn de basis waarop we oordelen. Het zijn de voorwaarden waaraan een (leerling)prestatie dient te voldoen om als succesvol te worden bestempeld. De criteria duiden welke de belangrijkste geachte leeruitkomsten zijn en welke kennis en vaardigheden leerlingen moeten kunnen demonstreren. Ze geven aan wat leerlingen moeten doen en welke aspecten van een prestatie in de beoordeling van tel zullen zijn.

Doelmatigheid. Doelmatigheid kunnen we zeer kernachtig omschrijven: iets is doelmatig als het geschikt is voor het doel waarvoor het gemaakt is. Het verzamelen en beoordelen van evaluatiegegevens verloopt met andere woorden doelmatig indien beantwoord wordt aan het doel waarvoor de gegevensverzameling ontworpen werd. Hiertoe worden de volgende drie doelmatigheidseisen voorop gesteld: *validiteit*, *betrouwbaarheid* en *efficiëntie*.

Evaluatie. Evaluatie is het verzamelen en beoordelen van gegevens over de (aan de doelstellingen van het onderwijs gerelateerde) prestaties van leerlingen.

Evaluatielijsten. Evaluatielijsten zijn instrumenten om leerlingprestaties te beoordelen en sluiten nauw aan bij de rubrieken en de criteria ontwikkeld werden voor de vooropgestelde evaluatietaak. De meeste evaluatielijsten bevatten een aantal uitspraken over kwaliteitsaspecten van een leerlingprestatie waarover de beoordelaar met behulp van een schaal een beoordeling uitspreekt.

Formatieve evaluatie. Bij formatieve evaluatie is het belang van de informatie vooral gelegen in de mogelijkheden tot feedback. Met dit evaluatietype wordt een tussentijdse vorm van evaluatie bedoeld, die de basis vormt voor de optimalisering en bijsturing van het leerproces van de leerling én het onderwijzen van de leerkracht.

Logboeken. Logboeken bestaan uit een continue weergave van activiteiten, ervaringen en reflecties van leerlingen. Het is als het ware een weergave van de innerlijke dialoog van de leerling. Het voornaamste doel van een logboek is de leerlingen te stimuleren om actief na te denken over hun eigen leerproces en om hen alzo hun eigen leerproces in handen te laten nemen. Een logboek wil leerlingen de kans geven gedachten, impressies en gevoelens naar boven te brengen, die anders geen aandacht dreigen te krijgen.

Objectiviteit. Met objectiviteit wordt benadrukt dat elke leerling recht heeft op een rechtvaardige en gelijke evaluatie. Factoren die niet relevant zijn voor het te evalueren leerlingenkenmerk mogen geen enkele invloed hebben op de evaluatieresultaten van de geëvalueerde. Het mag bijvoorbeeld geen verschil maken wie instaat voor de afname, correctie en waardering van de evaluatie.

Peerevaluatie. Een 'peer' is gewoonlijk een leeftijdsgenoot of iemand die in hetzelfde leerjaar zit. Leerlingen in dezelfde klas zijn dus mekaars peers. Peerevaluatie kan dan ook eenvoudig gedefinieerd worden als het evalueren door leerlingen (peers) van elkaars werk.

Portfolio's. Een portfolio is een doelgerichte verzameling van het werk van een leerling waarin diens inspanningen, vooruitgang en verwezenlijkingen aangetoond worden. Deze verzameling moet de deelname van de leerlingen omvatten in het selecteren van de inhoud van de portfolio, in het selecteren van criteria voor selectie van de inhoud en in het selecteren van criteria ter beoordeling van hun werk. Een portfolio omvat zowel duidelijke bewijzen van de reflectie van de leerling over diens eigen prestaties als de bemerkingen van de leerkracht op het werk en de reflectie van de leerling.

Rubriek. Een rubriek is een geschreven geheel van richtlijnen dat het mogelijk maakt prestaties en producten van verschillende kwaliteiten te onderscheiden. Een rubriek omvat beschrijvingen die aangeven waar men moet naar kijken bij elk kwaliteitsniveau. Ze maken duidelijk wat een bepaald prestatieniveau kenmerkt. Op die manier maken rubrieken duidelijk welke aspecten van de prestatie of het product het meest belangrijk zijn.

Standaard. Een standaard geeft aan welke criteria van belang zijn bij het beoordelen van een leerlingprestatie. Een standaard wordt dus geconcretiseerd aan de hand van criteria. Aan de hand van beschrijvingen en voorbeelden wordt in een standaard aangegeven hoe goed de prestatie van een leerling moet zijn om als excellent, voldoende of onvoldoende te worden bestempeld.

Summatieve evaluatie. Bij summatieve evaluatie wil men een eindoordeel uitspreken over de prestaties van een bepaalde leerling. Deze vorm van evaluatie wil dus als het ware een samenvatting geven van wat de leerling reeds bereikt heeft. Summatieve evaluatie is gericht op resultaatbepaling.

Vaardigheidsevaluatie. Vaardigheidsevaluatie is een evaluatievorm waarbij van de leerlingen verwacht wordt dat ze een antwoord construeren (in tegenstelling tot het kiezen van een antwoord uit een reeks mogelijkheden). Typisch aan vaardigheidsevaluatie is dat de leerlingen gevraagd worden om 1. bepaalde vaardigheden te demonstreren door iets te doen dat door de evaluatoren wordt geobserveerd op de moment dat het zich voordoet (een presentatie geven, een stopcontact aansluiten, een spel voor bejaarden begeleiden) of om 2. een taak te vervullen die resulteert in het creëren van een product (dit kan variëren van een kort geschreven antwoord tot een heuse tentoonstelling). Er wordt geëvalueerd of leerlingen de geleerde vaardigheden en competenties succesvol kunnen aanwenden.

Validiteit. De validiteit van een evaluatiemethode doelt op de eigenschap dat een evaluatiemethode al dan niet meet wat de gebruiker ervan bedoeld heeft ermee te meten. Anders geformuleerd: men moet de juiste conclusies kunnen trekken uit de evaluatieresultaten. Er zijn verschillende vormen van validiteit. Om gestalte te geven aan didactische evaluatie zijn de volgende vier het meest relevant: inhoudsvaliditeit, begripsvaliditeit, predictieve validiteit en levensechtheid.

Zelfevaluatie. Zelfevaluatie verwijst naar het evalueren door leerlingen van hun eigen kennis, vaardigheden, producten, leerprocessen en attitudes. De leerling richt zich systematisch op de eigen prestaties, meestal met als doel de toekomstige prestaties te verbeteren. De leerling krijgt hiertoe criteria, rubrieken en standaarden ter beschikking of ontwikkelt deze zelf. Zelfevaluatie kan ook bestaan uit het bekritisieren of simpelweg beschrijven van de eigen prestaties.

Overzicht van figuren en denkoefeningen

Ten geleide

Figuur 0.1: Vormen van evaluatie	7
Figuur 0.2: Structuur van het handboek	9
Figuur 0.3: Onduidelijkheid over begrippen	13

Deel 1: Stilstaan bij evaluatie

Figuur 1.1: Structuur van deel 1	19
Figuur 1.2: Rollen van de lerende	27
Figuur 1.3: De traditionele relatie tussen instructie en evaluatie	31
Figuur 1.4: Voorbeelden van multiple-choice- en waar-vals vragen	31
Figuur 1.5: Van een toetscultuur naar een assessmentcultuur	34
Figuur 1.6: De evaluatiefuncties binnen een assessmentcultuur	35
Figuur 1.7: De integratie van instructie, planning en evaluatie	37
Figuur 1.8: Illustratie van vakgebonden en vakoverschrijdende eindtermen	44
} Even stilstaan... 1.1: Samenhang tussen doelen, curriculum en evaluatie	21
} Even stilstaan... 1.2: Kritieken op de huidige onderwijspraktijk	26
} Even stilstaan... 1.3: Sporen van een traditionele evaluatiebenadering	33
} Even stilstaan... 1.4: Evaluatiefuncties in de huidige onderwijscontext	35
} Even stilstaan... 1.5: Traditionele en hedendaagse onderwijsbenaderingen	40

Deel 2: Op zoek naar goede evaluatie

Figuur 2.1: Structuur van deel 2	60
Figuur 2.2: Een taxonomie van onderwijsdoelstellingen	67
Figuur 2.3: De validiteit van een evaluatie bewaken	76
Figuur 2.4: Het verschil en het verband tussen betrouwbaarheid en validiteit	77
Figuur 2.5: Bronnen van meetfouten bij het evalueren	79
Figuur 2.6: Verschillende aspecten van betrouwbaarheid	79
Figuur 2.7: Kritische vragen om de doelmatigheid van een evaluatie te beoordelen	81
Figuur 2.8: Kritische vragen om de billijkheid van een evaluatie te beoordelen	83
Figuur 2.9: Continua richting alternatieve evaluatievormen	91
Figuur 2.10: Indicaties van de integratie van onderwijzen en evalueren	94
Figuur 2.11: Kenmerken van goede alternatieve evaluatie	104

Even stilstaan... 2.1: De validiteit van een evaluatie-instrument	73
Even stilstaan... 2.2: Niet het instrument is al dan niet valide, wel de conclusies	73
Even stilstaan... 2.3: De verschillende soorten van validiteit van een evaluatie	74
Even stilstaan... 2.4: De betrouwbaarheid van een evaluatie-instrument	77
Even stilstaan... 2.5: Aspecten van objectiviteit	82
Even stilstaan... 2.6: Hoe kwaliteitsvol is de evaluatie van de evaluatie?	89
Even stilstaan... 2.7: De evaluatie aanpassen aan kenmerken van de leerling	93
Even stilstaan... 2.8: Kritieken op geïsoleerde instructie en evaluatie	98
Even stilstaan... 2.9: De betrokkenheid van leerlingen	100
Even stilstaan... 2.10: Overlappende evaluatiekenmerken	105

Deel 3: Aandacht voor het evaluatieproces

Figuur 3.1: Structuur van het evaluatieproces	113
Figuur 3.2: Vragen bij het plannen van een evaluatie	118
Figuur 3.3: Kenmerken van goede evaluatie	123
Figuur 3.4: Ondersteunende vragen bij het beschrijven van evaluatietaken	126
Figuur 3.5: Een selectie van criteria	128
Figuur 3.6: Stappen bij het ontwikkelen van evaluatiecriteria	134
Figuur 3.7: Mogelijke criteria voor de dimensies van 'goed schrijven'	137
Figuur 3.8: Een rubriek ter illustratie	138
Figuur 3.9: Voorbeeld van een taakgebonden rubriek	140
Figuur 3.10: Voorbeeld van een algemene niet-taakgebonden rubriek	141
Figuur 3.11: Een analytische rubriek voor het geven van een mondelinge presentatie	142
Figuur 3.12: Een holistische rubriek voor het geven van een mondelinge presentatie	143
Figuur 3.13: Gebruik en voor- en nadelen van de verschillende rubriektypes	144
Figuur 3.14: Checklist ter beoordeling van criteria	145
Figuur 3.15: Rubriek voor het evalueren van rubrieken	148
Figuur 3.16: Scoringsrichtlijnen	150
Figuur 3.17: Een evaluatielijst voor het beoordelen van groepswork	152
Figuur 3.18: Differentieerbare evaluatielijst	153
Figuur 3.19: Een gecombineerde 10-puntenschaal	157
Figuur 3.20: Voorbeelden van viergradenschalen	160
Figuur 3.21: Voorbeeld van de '6-punten, twee-beslissingen'-schaal	161
Figuur 3.22: De betrouwbaarheid van een beoordelingsproces beoordelen	164
Figuur 3.23: Een voorbeeld van een narratief rapport	168
Even stilstaan... 3.1: Inventariseren van doelen	114
Even stilstaan... 3.2: De functies van evaluatie in uw onderwijspraktijk	117
Even stilstaan... 3.3: Invulling van kwaliteitseisen aan evaluatietaken	123
Even stilstaan... 3.4: Het belang van evaluatiecriteria	127
Even stilstaan... 3.5: De beschikbaarheid van evaluatiecriteria voor leerlingen	131

Even stilstaan... 3.6: Criteria clusteren tot leeruitkomsten	138
Even stilstaan... 3.7: Holistisch of analytisch?	144
Even stilstaan... 3.8: Het hardnekkig voortbestaan van puntensystemen	161

Deel 4: Alternatieve evaluatievormen geïllustreerd

Figuur 4.1: Continua richting alternatieve evaluatievormen	176
Figuur 4.2: Vaardigheidsevaluatie als alternatieve evaluatievorm	178
Figuur 4.3: Voorbeelden van een vaardigheidstaak	178
Figuur 4.4: Een aantal karakteristieken van vaardigheidstaken	183
Figuur 4.5: Voorbeelden van vaardigheidstaken	186
Figuur 4.6: Evaluatietaken schrijfvaardigheden	187
Figuur 4.7: Rubriek voor het invullen van formulieren	187
Figuur 4.8: Evaluatietaken voor spreekvaardigheden	188
Figuur 4.9: Holistische rubriek voor spreekvaardigheid	188
Figuur 4.10: Een rekentaak	189
Figuur 4.11: Een holistische rubriek voor het oplossen van rekenproblemen	189
Figuur 4.12: Een evaluatielijst voor het oplossen van rekenproblemen	189
Figuur 4.13: Storingen in signalisatie-installaties opsporen	190
Figuur 4.14: Holistische rubriek voor het opsporen van fouten	190
Figuur 4.15: checklist opsporen van fouten in een signalisatie-installatie	191
Figuur 4.16: Rubriek omtrent het begrip afhankelijkheid	192
Figuur 4.17: Authentieke evaluatie als alternatieve evaluatievorm	193
Figuur 4.18: Een aantal karakteristieken van authentieke taken	196
Figuur 4.19: Komen tot authentieke evaluatietaken	197
Figuur 4.20: Bronnen voor het ontwikkelen van authentieke taken	198
Figuur 4.21: Taken die een beroep doen op schrijfvaardigheden	200
Figuur 4.22: Een analytische schrijfrubriek	200
Figuur 4.23: Authentieke evaluatietaken voor luister- en schrijfvaardigheden	201
Figuur 4.24: Een taakgebonden rubriek	201
Figuur 4.25: Een authentieke rekentaak	202
Figuur 4.26: Rubriek voor het aanpakken van rekenproblemen	203
Figuur 4.27: Een authentieke elektriciteitsoefening	203
Figuur 4.28: Een holistische rubriek voor stage-evaluatie in de kleuterschool	204
Figuur 4.29: Evaluatielijst voor stage-evaluatie in de kleuterschool	205
Figuur 4.30: Leerlingenbetrokkenheid als kenmerk van alternatieve evaluatievormen	207
Figuur 4.31: Een continuüm van leerlingenbetrokkenheid	208
Figuur 4.32: Voorbeeld van een zelfevaluatie-instrument	210
Figuur 4.33: Zelfevaluatiecriteria	220
Figuur 4.34: Een instrument voor zelf-, peer- en co-evaluatie	221
Figuur 4.35: Zelfevaluatielijsten geïllustreerd	224
Figuur 4.36: Attitudeschaal 'relaties met medeleerlingen'	227

Figuur 4.37: Een toelichting voor leerlingen	228
Figuur 4.38: Een stagelogboek	229
Figuur 4.39: Een voorgestructureerd logboek	229
Figuur 4.40: Een voorbeeld van een logboekbijdrage	231
Figuur 4.41: Navertellen, relateren en reflecteren	232
Figuur 4.42: Leidraad voor een peerinterview	232
Figuur 4.43: Algemene voorbeelden van leerlingenbetrokkenheid	234
Figuur 4.44: Leerlingen betrekken bij het ontwikkelen van criteria	236
Figuur 4.45: Groepseigen logboek	237
Figuur 4.46: Individueel logboek	237
Figuur 4.47: Evaluatielijst 'samen in groep werken'	238
Figuur 4.48: Berekenen van gemiddelde en mediaan	238
Figuur 4.49: Evaluatielijst voor reflectie	239
Figuur 4.50: Holistische rubriek 'Goed reflecteren'	240
Figuur 4.51: Attitudes bij wiskunde in de eerste graad secundair onderwijs	240
Figuur 4.52: Attitudeschaal 'houding tegenover wiskunde'	241
Figuur 4.53: rubriek voor de attitude 'efficiëntie en zoeken naar oplossingsmethoden'	242
Figuur 4.54: Zelfevaluatielijst 'toepassen van veiligheidsvoorschriften'	243
Figuur 4.55: Activiteitenlogboek	244
Figuur 4.56: Learner report - leerlogboek	245
Figuur 4.57: Richtvragen voor een persoonlijk logboek	245
Figuur 4.58: Evaluatielijst 'stage bejaardentehuis'	246
Figuur 4.59: De alternatieve kenmerken van portfoliogebruik	247
Figuur 4.60: Een mogelijk gebruik van een portfolio	248
Figuur 4.61: Andere portfoliodoelen	252
Figuur 4.62: Voorbeelden van categorieën in een wiskundeportfolio	256
Figuur 4.63: Voorbeelden van de inhoud van portfolio's	257
Figuur 4.64: Portfolio-criteria	258
Figuur 4.65: Hoe succesvol leren selecteren?	260
Figuur 4.66: Reflectie over gans de portfolio	260
Figuur 4.67: Reflectie bij een specifiek werkstuk	261
Figuur 4.68: De verschillende componenten van een portfolio	265
Figuur 4.69: Holistische rubriek voor het beoordelen van een portfolio	265
Figuur 4.70: Analytische rubriek voor het beoordelen van een portfolio	266
Even stilstaan... 4.1: Constructiegerichtheid in Vlaamse scholen	180
Even stilstaan... 4.2: Een continuüm van constructiegerichtheid	182
Even stilstaan... 4.3: De geschiktheid van vroegere evaluatieinstrumenten	235
Even stilstaan... 4.4: Wat wil men bereiken met reflectie?	252

Deel 5: Aanvangen met alternatieve evaluatievormen

Figuur 5.1: Potentiële valkuilen	273
Figuur 5.2: De rol van beleidsmakers	279
Figuur 5.3: Aandachtspunten voor leerkrachten	282
Even stilstaan... 5.1: Samenhang tussen onderwijsdoelen, instructie en evaluatie	274
Even stilstaan... 5.2: De noodzaak van tijdsinvestering	277
Even stilstaan... 5.3: De scholing van leerkrachten	280
Even stilstaan... 5.4: Tijdsplanning van een invoeringsproces	282
Even stilstaan... 5.5: Uw eigen nascholingsbehoefte	283

Literatuurlijst

Deze literatuurlijst en een overzicht van websites die handelen over alternatieve evaluatievormen kunnen tevens worden geraadpleegd op www.edubron.be.

- Abercrombie, M.L.J. (1991). Changing basic assumptions about teaching and learning, in: D. Boud (ed.). *Developing autonomy in learning*. London: Kogan Page.
- Airasian, P.W. (1994). *Classroom assessment*. New York: McGraw-Hill.
- Arrasmith, D.G., Kushman, W. & Olson, N. (1996). Using Performance Assessment Information for Improvement. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Arter, J. & V. Spandel (1991). *Using portfolios of student work in instruction and assessment*. Portland: Northwest Regional Educational Laboratory.
- Arter, J.A. & Bond, L. (1996). Why is assessment changing?. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Arter, J.A. (1996b). Performance Criteria: the heart of the matter. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Arter, J. (2000). *Rubrics, scoring guides, and performance criteria: Classroom tools for assessing and improving student learning*. Paper presented at the annual conference of the American Educational Research Association, New Orleans.
- ASCD - Association for Supervision and Curriculum Development (1996). *Performance-based learning and assessment*. ASCD, Virginia.
- Aschbacher, P. R. (1993). *Issues in innovative assessment for classroom practice: Barriers and facilitators*. (CSE Tech. Rep. No. 359). Los Angeles: University of California, Center for Research on Evaluation, Standards and Student Testing (CRESST).
- Aschbacher, P.R., Koency, G., and Schacter, J. (1995). *Alternative Assessment Guidebook*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Baker, E., Herman, J., Gearhart, M. (1994). Assessing Writing Portfolios: Issues in the Validity and Meaning of Scores. *Educational Assessment*, 1 (3), pp.201-224.
- Baker, E. L., Linn, R. L & Herman, J. L. (1996). *CRESST: A continuing mission to improve educational assessment*. Evaluation comment. Los Angeles: University of California, Center for the Study of Evaluation & The National Center for Research on Evaluation, Standards, and Student Testing.

- Bartels, E. (1998). Developing reflective learners – Students self-Assessment assessment learning. *Journal of Professional Nursing*, vol. 14, No. 3, 1998, p. 135.
- Birenbaum, M. (1995). Assessment 2000: Towards a pluralistic approach to assessment. In: Birenbaum, M. & Dochy, F.R.C. (eds) (1995). *Alternatives in assessments of achievements, learning processes and prior knowledge*. Dordrecht: Kluwer Academic Publishers.
- Black, P. & Wiliam, D. (1998). Assessment and Classroom learning. *Assessment in Education*, vol. 5, No. 1, 1998.
- Blythe, T., Allen, D. & Powell B.S. (1999). *Looking Together at Student Work*. NY: Teachers College Press (1999).
- Bond, L.A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation*, 5(2).
- Broadfoot, P.M. (1996). *Education, assessment and society*. A sociological analysis. Buckingham: Open University Press.
- Brown, S., Race, P. and Bull, J. (1999). *Computer Assisted Assessment*. Kogan Page.
- Brualdi, A. (1998). Teacher comments on report cards. *Practical Assessment, Research & Evaluation*, 6(5).
- Brualdi, A. (2000). Implementing Performance Assessment in the Classroom. *Practical Assessment, Research & Evaluation*, 6(2).
- Burke, K., Fogerty, R. & Belgrad, S. (1994). *The Portfolio Connection: The Mindful School Series*. IRI/Skylight Publishing.
- Callison, D., (1998). Authentic Assessment. *School library Media Activities Monthly*, 14, nr. 5.
- Cardinet, J. (1995) Geschiedenis van de onderwijsevaluatie vanaf de oorsprong tot morgen. *Impuls voor onderwijsbegeleiding*, 4.
- Carnevale, A., Gainer, L. J., & Meltzer, A. S. (1996). *Workplace Basics: The Skills Employers Want*. Alexandria, VA: US Department of Labor Employment and Training Administration and American Society for Training and Development.
- Chapman, C. (1990). Authentic writing assessment. *Practical Assessment, Research & Evaluation*, 2(7).
- Cheng, W. & Warren M. (1997). 'Having second thoughts: student perceptions before and after a peer assessment exercise', *Studies in Higher Education* 22(2): 233-9.
- Clemens, J., Alberts, R., Hamelink, K. & Noordzij, P. (1999). Zelfevaluatie bij het leren: Ideeën en instrumenten. In Cito: *Toetsdeskundigheidsontwikkeling*. Cito, Arnhem.
- Conley, D.T. (1996). Assessment. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Creemers, B.P.M. (1991). *Effectieve instructie. Een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's-Gravenhage: : Instituut voor Onderzoek van het Onderwijs SVO.

- Darling-Hammond, L., Ancess, J. & Falk, B. (1995). *Authentic Assessment in Action: Studies of Schools and Students at Work*. New York: Teachers College Press.
- De Boer, E., de Jong, J. & van Warmerdam, P. (1998). *Handboek Vaardigheden voor de Tweede Fase van het Voortgezet Onderwijs*. Loenen aan de Vecht: Edumedia.
- De Corte, E. (1990). Towards powerful learning environments. *Journal of Psychology of Education*, 5 (1), 5-19.
- De Groot, A.D. & Wijnen, W.H. (1983). *Vijven en zessen: cijfers en beslissingen. Het selectieproces in ons onderwijs*. Groningen, Wolters-Noordhoff.
- De Neve, H., & Janssen, P.J. (1992). *Succesvol examineren in het hoger onderwijs*. Leuven-Amersfoort: Acco.
- De Rijke, T.R. & Beemer, C.F.H. (1995). Zelf-assessment. In: F. Dochy en T. de Rijke (eds), *Assessment centers: Nieuwe toepassingen in opleiding, onderwijs en HRM*. Utrecht, Lemma, p 75-91.
- Dietel, R. (1993). *What works in performance assessment?* Proceedings of the 1992 CRESST Conference. CRESST, UCLA.
- Dochy, F. & Segers, M. (1999). *Innovatieve toetsvormen als gevolg van constructiegericht onderwijs: op weg naar een assessmentcultuur*. Paper gepresenteerd op het UA-seminarie onderwijskundige professionalisering, 28.04.99.
- Dochy, F. (1999b). *Een structureel beleid voor toetsing en assessment in constructiegericht onderwijs (CGO): Kern voor het slagen van een universitaire onderwijsvernieuwing*. Limburg: LUC.
- Dochy, F., & De Rijke, T.R. (1995). *Assessment centers: nieuwe toepassingen in onderwijs, opleidingen en HRM*. Utrecht: Lemma.
- Dochy, F., Dierick, S. (2000). Assessment in Constructiegericht Onderwijs en Probleemgestuurd Onderwijs: een noodzakelijke voorwaarde voor universitaire onderwijsinnovatie. *Tijdschrift voor Onderwijsrecht en -beleid*, 4, 290-306.
- DVO (Dienst voor Onderwijsontwikkeling) (1997). *Wegwijs in de eindtermen en de ontwikkelingsdoelen eerste graad secundair onderwijs*. Brussel: Ministerie van de Vlaamse gemeenschap, AO-AID.
- DVO (Dienst voor Onderwijsontwikkeling) (1998). *Kerdoelen en eindtermen. Resultaten van een vergelijkende analyse van de Nederlandse basisvorming en de Vlaamse eindtermen eerste graad secundair onderwijs*. DVO, Brussel.
- Educational Resources Information Center. ERIC Definitions for Concepts Related to Alternative Assessment (no date). Online beschikbaar: www.ericae.net.
- Enger, S.K. & Yager, R.E. (2001). *Assessing student understanding in science: a standards-based K-12 handbook*. Thousands Oaks : Corwin Press.
- Erwin, T. D. (1991). *Assessing student learning and development: A guide to the principles, goals, and methods of determining college outcomes*. San Francisco: Jossey-Bass.
- Fullan, M.G. (1996). Overcoming barriers to educational change. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Gearhart, M. & Herman, J.L. (1995). *Portfolio Assessment: Whose Work Is It? Issues in the Use of Classroom Assignments for Accountability*. Los Angeles: Center for the Study.
- George, P. (1995). *What Is Portfolio Assessment Really and How Can I Use It in My Classroom?* Gainesville, FL: Teacher Education Resources.
- Gipps, C.V. (1994). *Beyond testing. Towards a theory of educational assessment*. London: The Falmer Press.
- Glaser, R. (1990). Toward new models for assessment, *International Journal of Educational Research*, 14, 375-483.
- Gredler, M.E. (1999). *Classroom Assessment and learning*. Longman: New York.
- Harris, D. & Bell, C. (1990). *Evaluating and assessing for learning*. New York: Nichols Publishing Company.
- Harris, D. and C. Bell. (1994). *Evaluating and Assessing for Learning, revised edition*. Londen: Kogan Page.
- Harris, M. & McCann, P. (1994). *Assessment*. Oxford: Heinemann Publishers.
- Harskamp, E.G. (1988). *Rekenmethoden op de proef gesteld*. Groningen: RION, Instituut voor Onderwijsonderzoek. diss. RUG 7-7-1988.
- Herman, J. (1996). Technical quality matters. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Herman, J., Aschbacher, P.R. & Winters, L. (1996b). Setting Criteria. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Herman, J.L. & Knuth, R.A. (1991). *What Does Research Say About Assessment?* NCREL, Oak Brook.
- Herman, J.L., Morris, L.L., & Fitz-Gibbon, C.T. (1987). *Evaluator's Handbook*. (Second edition.). Newbury Park, Sage Publications.
- Herman, L., Aschbacher, R. & Winters, L. (1992). *A practical Guide to Alternative Assessment*. Association for Supervision and Curriculum Development, Alexandria Virginia.
- Hibbard, K.M. (1996). Self-Assessment Using Performance Task Assessment Lists. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Hofstee, W.K.B. (1999). *Principes van Beoordeling. Methodiek en ethiek van selectie, examineren en evaluatie*. Lisse, Swets en Zeitlinger.

- Imbrecht, I. & Van Petegem, P. (2000). *Leertheorieën en onderwijspsychologische theorieën*. Openleerpakket voor het Project 'Begeleid Zelfstandig leren'. Universiteit Antwerpen.
- Inspectie van het onderwijs (2001). *Onderwijsspiegel: Verslag over de toestand van het onderwijs (schooljaar 1999-2000)*. Ministerie van de Vlaamse Gemeenschap, Brussel.
- Janssen, P.J. (1999). Als examiner slagen in eigen examineren. Aspecten van examineerexpertise. *Tijdschrift voor Onderwijsrecht en onderwijsbeleid*, 6, 444-459.
- Janssens, S., Verschaffel, L., De Corte, E., Lowyck, J., Struyf, E., Van Damme, J & Vandenberghe, R. (2000). *Didactiek in beweging*. Wolters, Plantyn.
- Kerka, S. (1995). *Techniques for Authentic Assessment*. ERIC - Practice Application Brief - adult, career, and vocational education. Online beschikbaar: www.ericacve.org
- Klasse voor Leerkrachten (1997). *Eindtermen: Het kompas voor de leraar*. 75, mei 1997, p. 30-31.
- Klasse voor Leerkrachten (1998). *Scholen zonder examens*. 86, juni 1998, p. 32-33
- Koretz, D., Stecher, B., Klein, S., McCaffrey, D. & Deibert, E. (1993). *Can Portfolios Assess Student Performance and Influence Instruction? The 1991-92 Vermont Experience*. RAND Institute on Education and Training
- Kulieke, M., Bakker, J., Collins, C., Fennimore, T., Fine, C., Herman, J., Jones, B.F., Raack, L. & Tinzmann, M.B. (1990). *Why Should Assessment Be Based on a Vision of Learning?* NCREL, Oak Brook.
- Lazaro, W. L. (1996). Empowering students with instructional rubrics. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Longhurst, N. & Norton, L.S. (1997). Self-assessment in coursework essays. *Studies in Educational Evaluation*, 23, 4, 319 - 330.
- Lowyck, J. & Verloop, N. (1995). *Onderwijskunde. Een kennisbasis voor professionals*. Wolters Leuven.
- Mabry, L. (1999). *Portfolios plus: a critical guide to alternative assessment*. Thousand Oaks, California: Corwin Press
- Madaus, G. & Kellaghan, T. (1992). Curriculum, evaluation and assessment. In: P.W. Jackson (Red.), *Handbook of research on curriculum*. A project of the AERA, p. 119-154. NY: Macmillan Publishing Company.
- Maeroff, G. I. (1991). Assessing Alternative Assessment. *Phi Delta Kappan*, 73, 273-281.
- McLean, L.D. (1990). Time to replace the classroom test with authentic measurement. *Alberta Journal of Educational Research*, 36, 78-84.
- McMillan, J. H. (2001). *Essential assessment concepts for teachers and administrators*. Thousand Oaks, CA: Corwin Publishing Company.
- Miller, A.H., Imrie, B.W. & Cox, K. (1998). *Student Assessment in Higher Education. A handbook for assessing performance*. London: Kogan Page Limited.
- Ministerie van de Vlaamse Gemeenschap (1996a). *Decreet tot bekrachtiging van de eindtermen en de ontwikkelingsdoelen van de eerste graad van het gewoon secundair onderwijs*. 24 juli 1996: Brussel.
- Ministerie van de Vlaamse Gemeenschap (1999e). *Besluit van de Vlaamse regering betreffende de basiscompetenties van de leraren (B.S. 16-2-1999) – 29 september 1998*. Brussel.
- Ministerie van de Vlaamse Gemeenschap (2000). *Besluit van de Vlaamse Regering tot de vaststelling van de eindtermen van de tweede en derde graad van het gewoon secundair onderwijs (23 juni 2000 - B.S. 29-11-2000)*. Brussel.
- Mitchell, R. (1992). *Testing for learning. How new approaches to evaluation can improve American Schools*. New York: Free Press.
- MMS (Miami Museum of Science), (2001). *Forms of alternative assessment*. Online beschikbaar: www.miamisci.org
- Morris, L.L., Fitz-Gibbon, C.T., and Freeman, Marie E. (1987). *How to Communicate Evaluation Findings*. Newbury Park, Sage Publications.
- Moskal, B. (2000). Scoring rubrics: what, when and how?. *Practical Assessment, Research & Evaluation*, 7(3).
- Moskal, B. M. & Jon A. L. (2000). Scoring rubric development: validity and reliability. In *Practical Assessment, Research & Evaluation*, 7(10).
- Moya, S.S. & O'Malley, J.M. (1994). A portfolio assessment model for ESL. *The Journal of Educational Issues of Language Minority Students*, v13 p. 13-36, Spring 1994.
- Neil, D.M. & N.J. Medina (1989). Standardized testing: harmful to educational health. *Phi Delta Kappan*, 70, 688-697.
- Niss, M. (1993). *Investigations into Assessment in Mathematics Education: An ICMI Study*. Dordrecht: Kluwer.
- Nixon, N.J. (1990). Assessment of experience-based learning., in: C. Bell & D. Harris (eds.). *Assessment and evaluation*. World Yearbook of Education 1990. London: Kogan Page.
- O'neil, J. (1992). Putting performance assessment to the test. *Educational leadership*, 49, p. 14-19.
- OECD (1998). *Making the curriculum work*. Paris: OECD/CERI.
- Office of Research (1993). Student Portfolios: Classroom Uses. *Education Research Consumer Guide*, Nr 8., U.S. Department of Education.
- P.L.U.S (pedagogical language usage server) (2001). Dictionary of education. <http://macweb.soe.umich.edu/plus/index.html>

- Picus, L. and A. Tralli (1998) *Alternative Assessment Programs: What are the True Costs? An Analysis of the Total Costs of Assessment in Kentucky and Vermont*. Los Angeles, CA: National Center for Research on Evaluation, Standards and Student Testing.
- Popham, W.J. (1995). *Classroom assessment: What teachers need to know*. Boston: Allyn and Bacon.
- Putten, C.M. van (1987). *Leerlingen van het individueel beroepsonderwijs nader beschouwd. Een vergelijkend onderzoek naar cognitieve, sociaal-emotionele en achtergrondkenmerken van leerlingen bij de aanvang van het IBO, LBO en VSO/MLK*. diss. RUL, 21-10-1987. (SVO-project 0387).
- Roby, B. (1990). Assessment and the curriculum. In: Riding, R. & Butterfield, S. (Eds). *Assessment and examination in the secondary school*. London and NY: Routledge.
- Roeber, E. D. (1996). Guidelines for the development and management of performance assessments. *Practical Assessment, Research & Evaluation*, 5(7).
- Rolheiser, C., Bower, B., and Stevahn, L., (2000). *The Portfolio Organizer*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Rudman, H. C. (1989). Integrating testing with teaching. *Practical Assessment, Research & Evaluation*, 1(6).
- Scriven, M. (1969). An introduction to meta-evaluation. *Educational Product Report*, 2, 36-38.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 26(7), 4-14.
- SLI (Sint-Lutgardisinstituut Mol), (2002). Niet gepubliceerd stagebeoordelingsformulier.
- Sluijsmans, D., Dochy, F., & Moerkerke, G. (1999). Creating a Learning Environment by Using Self- Peer- and Co-assessment. *Learning Environments Research*, 1, 293-319.
- Soled, S.W. (Ed.) (1995). *Assessment testing and evaluation in teacher education*. Norwood, NJ: Ablex.
- Soudant, R. (2002). *Vademecum: De geïntegreerde proef in de Provinciale Secundaire School te Voeren*. Online beschikbaar: <http://www.ping.be/~rsoudantframes.html>
- Spandel, V. & Culham, R. (1996). Writing assessment. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Standaert, R. & Troch, F. (1998). *Leren en onderwijzen. Inleiding tot de algemene didactiek*. Leuven, Acco.
- Stiggins, R.J. (1996). Keeping Performance assessment in Perspective. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.

- Stiggins, R.J. (2001). *Student-Involved Classroom Assessment*, 3rd Ed. New Jersey: Prentice-hall.
- Struyf, E., Vandenberghe, R., & Lens, W. (1999). Recente ontwikkelingen in het denken over evaluatie. *Handboek Leerlingenbegeleiding*, 25, p. 179
- Struyf, E., Vandenberghe, R., & Lens, W. (1999b). Het formatieve en summatieve karakter van schriftelijke toetsen, *Handboek Leerlingenbegeleiding*, 25, 21-36.
- Struyf, E. (2000). *Evalueren: een leerkans voor leraren en leerlingen? Over het evaluatiepraktijk in de klas en het evaluatiebeleid op school*. Leuven: Universitaire pers.
- Thio. (1988) Toetstechnische Begrippenlijst. Cito, Arnhem.
- Tillema, H.H. (1998). Design and validity of a portfolio instrument for professional training, *Studies in Educational Evaluation*, 24(3), pp. 263-278.
- Tondeur, N. (1998). *Assessing Outcomes-based learning*. Free state Department of Education, Bloemfontein.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276.
- Torrance, Harry. (Ed.). (1995). *Evaluating Authentic Assessment*. Philadelphia, PN: Open University Press.
- van Gog T.A.J.M. & Heijmen I.L.V. (2001). *Portfolio nader bekeken. Een literatuuronderzoek naar (digitaal) portfolio*. Fontys.
- Verhoef, A.C. & Visser, L. (1994). *Omgangskunde 1*. Nijgh & Van Ditmar Educatief.
- Verhoeven, J., Warmoes, V., Devos, G., van Petegem, P., Bruylant, W., & Vanhoof, J. (2001). *Het evaluerend vermogen van scholen*. Niet gepubliceerd rapport bij het O.B.P.W.O.-Project 99.11 uitgevoerd in opdracht van het Departement Onderwijs.
- Verloop, N. & van de Schoot, F. (1995). Didactische evaluatie. In: Lowyck, J & Verloop, N. (Red.), *Onderwijskunde. Een kennisbasis voor professionals*. Leuven, Wolters.
- Verschaffel, L. (1995). Beïnvloeden van leerprocessen. In J.Lowyck & N. Verloop (red.), *Onderwijskunde. Een kennisbasis van professionals*. Groningen: Wolters-Noordhoff.
- Verschaffel, L., Vermunt, J.D. (eds.) (1998). *Het leren van leerlingen*. In: *Onderwijskundig Lexicon*. Editie III. Alphen aan den Rijn: Samsom.
- Waltman, K., Kahn, A. & Koency, G. (1998). *Alternative Approaches to Scoring: The Effects of Using Different Scoring Methods on the Validity of Scores From a Performance assessment*. CSE technical report 488. University of California: Los Angeles.
- WCER (1993). *Mathematics Performance Assessment Task Criteria*. Wisconsin Center for Educational Research.
- Webb, N. & Coxford, F. (1993). *Assessment in the Mathematics Classroom*. Reston: NCTM.

- Westerberg, T. (1996). What we've learned: reflections on implementing performance assessment. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, 2(2).
- Wiggins, G. (1994). Toward better report cards. *Educational Leadership*. 52 (2).
- Wiggins, G. (1996). What is a Rubric? A dialogue on design and use. In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Wiggins, G. P. (1998). *Educative assessment*. San Francisco: Jossey-Bass Publishers.
- Winograd, P., & Perkins, F. (1996). Authentic assessment in the classroom: Principles and practices . In: *Student Performance assessment in an Era of Restructuring*, Blum R. E. and Arter, J.A. (Editors), Alexandria, VA: Association for Supervision and Curriculum.
- Wolfe, W. & Miller, R. (1997). Barriers to the Implementation of Portfolio assessment in Secondary education. *Applied Measurement in Education*, 10 (3), 235-251.
- Wood, R. (1991). *Assessment and testing*. Cambridge: Cambridge University Press.
- Zariski, A. (1996). Student peer assessment in tertiary education: Promise, perils and practice. In Abbott, J. and Willcoxson, L. (Eds), *Teaching and Learning Within and Across Disciplines*, p189-200.